

STIC-Biotech/ChemLib

103061

From: Gupta, Anish
Sent: Thursday, September 04, 2003 10:59 AM
T : STIC-Biotech/ChemLib
Subject: search request

Search Request:

Application Serial No. 09/580893
Art Unit : 1654
Phone: 703-308-4001
Location: 9a13

Please search: Seq. ID 55-65

thank you

anish gupta

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/4/03
Date Completed: 9/15/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: 11
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 03
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 103061

TO: Anish Gupta
Location: CM1/9A13&11D13
Art Unit: 1645
Friday, September 05, 2003

Case Serial Number: 09/580893

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Gupta,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 20.5882 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893D-55

Perfect score: 21

Sequence: 1 VVPN 4

Scoring table: BLASTUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	4	23	AAE16720	Ligamentum nuchae
2	100.0	5	23	AAE16721	Ligamentum nuchae
3	100.0	5	23	AAE16723	Ligamentum nuchae
4	100.0	6	19	AAE16720	Rodent IL-1 delta
5	100.0	6	23	AAE16722	Ligamentum nuchae
6	100.0	6	23	AAE16724	Ligamentum nuchae
7	100.0	6	23	AAE16725	Ligamentum nuchae
8	100.0	6	23	AAE16728	Ligamentum nuchae
9	100.0	7	12	AAE11971	Dolastatin 4, Do1

10	21	100.0	7	23	AAE16726	Ligamentum nuchae
11	21	100.0	7	23	AAE16729	Ligamentum nuchae
12	21	100.0	8	23	AAE16727	Ligamentum nuchae
13	21	100.0	9	16	AAE16730	Ligamentum nuchae
14	21	100.0	9	16	AAE16732	MART-1 melanoma an
15	21	100.0	9	22	AAE16730	MART-1 antigen pep
16	21	100.0	11	22	AAE16730	Vascular dementia-
17	21	100.0	11	22	AAE16730	Vascular dementia-
18	21	100.0	11	22	AAE16730	Human API-60 trypt
19	21	100.0	11	22	AAE16730	DPI tryptic digest
20	21	100.0	11	22	AAE16730	DPI tryptic digest
21	21	100.0	11	22	AAE16730	DPI tryptic digest
22	21	100.0	11	22	AAE16730	DPI tryptic digest
23	21	100.0	11	22	AAE16730	Schizophrenia-asso
24	21	100.0	11	22	AAE16730	Schizophrenia-asso
25	21	100.0	11	22	AAE16730	Schizophrenia-asso
26	21	100.0	11	22	AAE16730	Schizophrenia-asso
27	21	100.0	11	22	AAE16730	Schizophrenia-asso
28	21	100.0	11	22	AAE16730	Schizophrenia-asso
29	21	100.0	11	22	AAE16730	Schizophrenia-asso
30	21	100.0	11	22	AAE16730	Schizophrenia-asso
31	21	100.0	11	22	AAE16730	Schizophrenia-asso
32	21	100.0	11	22	AAE16730	Schizophrenia-asso
33	21	100.0	11	22	AAE16730	Schizophrenia-asso
34	21	100.0	11	22	AAE16730	Schizophrenia-asso
35	21	100.0	11	22	AAE16730	Schizophrenia-asso
36	21	100.0	11	22	AAE16730	Schizophrenia-asso
37	21	100.0	11	22	AAE16730	Schizophrenia-asso
38	21	100.0	11	22	AAE16730	Schizophrenia-asso
39	21	100.0	11	22	AAE16730	Schizophrenia-asso
40	21	100.0	11	22	AAE16730	Schizophrenia-asso
41	21	100.0	11	22	AAE16730	Schizophrenia-asso
42	21	100.0	11	22	AAE16730	Schizophrenia-asso
43	21	100.0	11	22	AAE16730	Schizophrenia-asso
44	21	100.0	11	22	AAE16730	Schizophrenia-asso
45	21	100.0	11	22	AAE16730	Schizophrenia-asso

ALIGNMENTS

RESULT 1	
ID AAE16720	standard; peptide; 4 AA.
AAE16720:	
XX	
AC	09-APR-2002 (first entry)
XX	
DT	
XX	
DE	Ligamentum nuchae hydrolysed elastin peptide VVPN.
XX	
XX	
KW	Hydrolysed elastin peptide; HEP, therapy; cosmetic appearance; cardiact;
KW	vasotrophic; hypertension; skin disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypertensive; chronic obstructive pulmonary disease; antidiabetic;
KW	restenosis post-angioplasty; antidiabetic; antidiabetic; thrombolytic.
XX	
OS	Ligamentum nuchae.
OS	Synthetic.
PN	W0200191700-A2.
PN	
PD	06-DEC-2001.
XX	
XX	
PF	30-MAY-2001; 2001WO-US17384.
XX	
XX	
PR	30-MAY-2000; 2000US-0580110.
PR	30-MAY-2000; 2000US-0580156.
PR	30-MAY-2000; 2000US-0580893.
PR	30-MAY-2000; 2000US-0584001.
XX	
XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX	

PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP) VVPN.
 CC
 SQ Sequence 4 AA;
 OY
 Query Match 100.0%; Score 21; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 VVPN 4
 1111
 1 VVPN 4
 RESULT 2
 AAE16721
 ID AAE16721 standard; peptide; 5 AA.
 XX
 AC AAE16721;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #1.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 OS
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide

PT substances which mimic the sequences found in elastin
 XX
 DR Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 CC
 SQ Sequence 5 AA;
 OY
 Query Match 100.0%; Score 21; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 2 VVPN 5
 1111
 2 VVPN 5
 RESULT 3
 AAE16723
 ID AAE16723 standard; peptide; 5 AA.
 XX
 AC AAE16723;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #3.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 5
 FT /note="C-terminal amide"
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin

XX Claim 24; Page 21; 53pp; English.
 XX
 PS The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 XX
 SQ Sequence 5 AA:

Query Match 100.0%; Score 21; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 ||||
 Db 2 VVPN 5

RESULT 4
 AAM86290
 ID AAM86290 standard; peptide; 6 AA.
 XX
 AC AAM86290;

DT 19-FEB-1999 (first entry)

DE Rodent IL-1 delta polypeptide fragment (residues 43-48).

KM Interleukin; IL-1 delta; polyclonal antibody; IL-1 epsilon; cytokine;
 KM inflammatory response; immune system; diagnosis; agonist; antagonist;
 KM chemokine.

OS Mus sp.

PN W09847921-A1.

PD 29-OCT-1998.

PF 17-APR-1998; 98MO-US06879.

PR 06-AUG-1997; 97US-0055111.

PR 21-APR-1997; 97US-0837627.

PA (SCHE) SCHERING CORP.

PI Bazan JF, Hedrick JA, Kastelein RA, Sana TR;

XX WPI; 1998-609976/51.

PT Mammalian interleukin 1-delta and 1-epsilon - useful for, e.g.

XX regulating the immune system and inflammatory responses

PS Claim 1; Page 99; 113pp; English.

XX The invention relates to a recombinant polypeptide that specifically
 CC binds polyclonal antibodies (Abs) generated against a 12 consecutive
 CC amino acid segment of interleukin (IL)-1 delta or IL-1 epsilon. Agonists
 CC or antagonists of these IL polypeptides are used to regulate a cell
 CC involved in an inflammatory response. The IL-1 delta or IL-1 epsilon
 CC polypeptides and peptides are used to produce Abs and antigen-Abs
 CC complexes. The polypeptides, Abs and the corresponding nucleic acids

CC regulate development and/or the immune system, and can be used to
 CC diagnose and treat conditions associated with abnormal expression of IL.
 CC Agonists or antagonists of IL-1 delta or IL-1 epsilon polypeptides are
 CC used with agonists or antagonists of IL-1 alpha, IL-1RA, IL-1 beta, IL-1
 CC gamma, IL-2 and/or IL-12. The IL-1 delta or IL-1 epsilon polypeptides
 CC may be used as a soluble polypeptide or as a fusion protein with another
 CC cytokine or chemokine. Sequences AAM86287 to AAM86300 represent peptide
 CC fragments of a rodent interleukin (IL)-1 delta polypeptide.
 XX
 SQ Sequence 6 AA:

Query Match 100.0%; Score 21; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 ||||
 Db 3 VVPN 6

RESULT 5
 AAE16722
 ID AAE16722 standard; peptide; 6 AA.
 XX
 AC AAE16722;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #2.

KM Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KM vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KM hypotensive; chronic obstructive pulmonary disease; antihypertensive;
 KM restenosis post-angioplasty; antihypertensive; antihypertensive; thrombolytic.

OS Ligamentum nuchae.
 OS Synthetic.

PN W0200191700-A2.

PD 06-DEC-2001.

PF 30-MAY-2001; 2001MO-US17384.

PR 30-MAY-2000; 2000US-0580110.

PR 30-MAY-2000; 2000US-0580156.

PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.

PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.

PI Miltis TF, Sandberg LB, Jimenez F;

XX WPI; 2002-106259/14.

PT Composition for improving the elasticity of tissue comprises peptide

XX substances which mimic the sequences found in elastin

PS Claim 24; Page 21; 53pp; English.

XX The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use

CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

Sequence 6 AA;

Query Match 100.0%; Score 21; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPN 4
|
|
|
|
Db 3 VVPN 6

RESULT 6
AAE16724
ID AAE16724 standard; peptide; 6 AA.

AC AAE16724;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #4.

KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW restenosis; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.
XX
XX Ligamentum nuchae.
OS Synthetic.

FT Key Location/Qualifiers
FT Modified-site 6 /note="C-terminal amide"

PN WO200191700-A2.

PD 06-DEC-2001.

PF 30-MAY-2001; 2001WO-US17384.

PR 30-MAY-2000; 2000US-0580110.

PR 30-MAY-2000; 2000US-0580156.

PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.

(CONN-) CONNECTIVE TISSUE IMAGING LLC.

PI Mitts TF, Sandberg LB, Jimenez F;

DR WPI; 2002-106259/14.

PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
XX Claim 24; Page 21; 53pp; English.

XX The present invention relates to a composition for the treatment of
XX mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
XX their biological equivalents. The composition is useful not only for
XX improving the cosmetic appearance of skin, but for the treatment of
XX various skin disorders associated with reduced elastin formation and for
XX numerous other tissue disorders especially of blood vessels including
XX hypertension, coronary heart disease, arteriosclerosis, angina, coronary
XX thrombosis, chronic obstructive pulmonary disease and restenosis post-
XX angioplasty. They may also be useful in coating surgical appliances i.e.
XX stents. The use of peptides and peptide fragments which mimic those found
XX in elastin overcomes the problem of solubility of the dense cross-linked
XX structure of elastin itself, which renders it difficult to use
XX therapeutically. The peptides are found to improve elastin production in

CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

Sequence 6 AA;

Query Match 100.0%; Score 21; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPN 4
|
|
|
|
Db 3 VVPN 6

RESULT 7
AAE16725
ID AAE16725 standard; peptide; 6 AA.

AC AAE16725;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #5.

KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.
XX
XX Ligamentum nuchae.
OS Synthetic.

FT Key Location/Qualifiers
FT Modified-site 6 /note="C-terminal amide"

PN WO200191700-A2.

PD 06-DEC-2001.

PF 30-MAY-2001; 2001WO-US17384.

PR 30-MAY-2000; 2000US-0580110.

PR 30-MAY-2000; 2000US-0580156.

PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.

(CONN-) CONNECTIVE TISSUE IMAGING LLC.

PI Mitts TF, Sandberg LB, Jimenez F;

DR WPI; 2002-106259/14.

PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
XX Claim 24; Page 21; 53pp; English.

XX The present invention relates to a composition for the treatment of
XX mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
XX their biological equivalents. The composition is useful not only for
XX improving the cosmetic appearance of skin, but for the treatment of
XX various skin disorders associated with reduced elastin formation and for
XX numerous other tissue disorders especially of blood vessels including
XX hypertension, coronary heart disease, arteriosclerosis, angina, coronary
XX thrombosis, chronic obstructive pulmonary disease and restenosis post-
XX angioplasty. They may also be useful in coating surgical appliances i.e.
XX stents. The use of peptides and peptide fragments which mimic those found
XX in elastin overcomes the problem of solubility of the dense cross-linked
XX structure of elastin itself, which renders it difficult to use
XX therapeutically. The peptides are found to improve elastin production in
XX treated tissues. The present sequence is Ligamentum nuchae hydrolysed
XX elastin peptide (HEP), VVPN derived peptide.

Sequence 6 AA;

OY 1 VVPN 4
ID 1111
DB 3 VVPN 6

RESULT 10

ID AAE16726 standard; peptide; 7 AA.

AC AAE16726;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #6.

KM Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;

KM vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;

KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;

KM hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;

KM restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.

OS Ligamentum nuchae.

OS Synthetic.

PN WO200191700-A2.

PF 30-MAY-2001; 2001WO-US17384.

PR 30-MAY-2000; 2000US-0580110.

PR 30-MAY-2000; 2000US-0580156.

PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.

PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.

PI Mitis TF, Sandberg LB, Jimenez F;

DR WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide

PT substances which mimic the sequences found in elastin

XX Claim 24; Page 21; 53pp; English.

PS The present invention relates to a composition for the treatment of

CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or

CC their biological equivalents. The composition is useful not only for

CC improving the cosmetic appearance of skin, but for the treatment of

CC various skin disorders associated with reduced elastin formation and for

CC numerous other tissue disorders especially of blood vessels including

CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary

CC thrombosis, chronic obstructive pulmonary disease and restenosis post-

CC angioplasty. They may also be useful in coating surgical appliances i.e.

CC stents. The use of peptides and peptide fragments which mimic those found

CC in elastin overcomes the problem of solubility of the dense cross-linked

CC structure of elastin itself, which renders it difficult to use

CC therapeutically. The peptides are found to improve elastin production in

CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed

CC elastin peptide (HEP), VVPN derived peptide.

XX Sequence 7 AA;

SO Query Match 100.0%; Score 21; DB 23; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
AAE16729
ID AAE16729 standard; peptide; 7 AA.

AC AAE16729;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #9.

KM Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;

KM vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;

KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;

KM hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;

KM restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;

XX cyclic.

XX Ligamentum nuchae.

OS Synthetic.

PN WO200191700-A2.

PF 06-DEC-2001.

PR 30-MAY-2001; 2001WO-US17384.

PR 30-MAY-2000; 2000US-0580110.

PR 30-MAY-2000; 2000US-0580156.

PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.

PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.

PI Mitis TF, Sandberg LB, Jimenez F;

DR WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide

PT substances which mimic the sequences found in elastin

XX Claim 24; Page 21; 53pp; English.

PS The present invention relates to a composition for the treatment of

CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or

CC their biological equivalents. The composition is useful not only for

CC improving the cosmetic appearance of skin, but for the treatment of

CC various skin disorders associated with reduced elastin formation and for

CC numerous other tissue disorders especially of blood vessels including

CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary

CC thrombosis, chronic obstructive pulmonary disease and restenosis post-

CC angioplasty. They may also be useful in coating surgical appliances i.e.

CC stents. The use of peptides and peptide fragments which mimic those found

CC in elastin overcomes the problem of solubility of the dense cross-linked

CC structure of elastin itself, which renders it difficult to use

CC therapeutically. The peptides are found to improve elastin production in

CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed

CC elastin peptide (HEP), VVPN derived cyclic peptide.

XX Sequence 7 AA;

SO Query Match 100.0%; Score 21; DB 23; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 VVFN 6

RESULT 12

ID AAE16727 standard; peptide; 8 AA.

AAE16727;

09-APR-2002 (first entry)

Ligamentum nuchae hydrolysed elastin peptide, VVFN derive peptide #7.

Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant; KM vasotropic; hypertension; skin disorder; tissue disorder; dermatological; KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina; KM hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic; KM stenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.

Ligamentum nuchae.
Synthetic.

WO200191700-A2.

06-DEC-2001.

30-MAY-2001; 2001WO-US17384.

30-MAY-2000; 2000US-0580110.

30-MAY-2000; 2000US-0580156.

30-MAY-2000; 2000US-0580893.

30-MAY-2000; 2000US-0584001.

(CONN-) CONNECTIVE TISSUE IMAGING LLC.

Mills TF, Sandberg LB, Jimenez F;

WPI; 2002-106259/14.

Composition for improving the elasticity of tissue comprises peptide

substances which mimic the sequences found in elastin

Claim 24; Page 21; 53pp; English.

The present invention relates to a composition for the treatment of mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or their biological equivalents. The composition is useful not only for improving the cosmetic appearance of skin, but for the treatment of various skin disorders associated with reduced elastin formation and for numerous other tissue disorders especially of blood vessels including hypertension, coronary heart disease, arteriosclerosis, angina, coronary thrombosis, chronic obstructive pulmonary disease and restenosis post-angioplasty. They may also be useful in coating surgical appliances i.e. stents. The use of peptides and peptide fragments which mimic those found in elastin overcomes the problem of solubility of the dense cross-linked structure of elastin itself, which renders it difficult to use therapeutically. The peptides are found to improve elastin production in treated tissues. The present sequence is Ligamentum nuchae hydrolysed elastin peptide (HEP), VVFN derived peptide.

Sequence 8 AA;

Query Match 100.0%; Score 21; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

1 VVFN 4
4 VVFN 7

RESULT 13

AAE16730
ID AAE16730 standard; peptide; 8 AA.

AAE16730;

09-APR-2002 (first entry)

Ligamentum nuchae hydrolysed elastin peptide, VVFN derived peptide #10.

Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant; KM vasotropic; hypertension; skin disorder; tissue disorder; dermatological; KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina; KM hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic; KM stenosis post-angioplasty; antianginal; anticoagulant; thrombolytic; cyclic.

Ligamentum nuchae.
Synthetic.

WO200191700-A2.

06-DEC-2001.

30-MAY-2001; 2001WO-US17384.

30-MAY-2000; 2000US-0580110.

30-MAY-2000; 2000US-0580156.

30-MAY-2000; 2000US-0580893.

30-MAY-2000; 2000US-0584001.

(CONN-) CONNECTIVE TISSUE IMAGING LLC.

Mills TF, Sandberg LB, Jimenez F;

WPI; 2002-106259/14.

Composition for improving the elasticity of tissue comprises peptide

substances which mimic the sequences found in elastin

Claim 24; Page 21; 53pp; English.

The present invention relates to a composition for the treatment of mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or their biological equivalents. The composition is useful not only for improving the cosmetic appearance of skin, but for the treatment of various skin disorders associated with reduced elastin formation and for numerous other tissue disorders especially of blood vessels including hypertension, coronary heart disease, arteriosclerosis, angina, coronary thrombosis, chronic obstructive pulmonary disease and restenosis post-angioplasty. They may also be useful in coating surgical appliances i.e. stents. The use of peptides and peptide fragments which mimic those found in elastin overcomes the problem of solubility of the dense cross-linked structure of elastin itself, which renders it difficult to use therapeutically. The peptides are found to improve elastin production in treated tissues. The present sequence is Ligamentum nuchae hydrolysed elastin peptide (HEP), VVFN derived cyclic peptide.

Sequence 8 AA;

Query Match 100.0%; Score 21; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

1 VVFN 4
4 VVFN 7

Db 4 VVPN 7

RESULT 14

ID AAR84772 standard; Peptide: 9 AA.

XX AAR84772;

AC AAR84772;

XX 25-APR-1996 (first entry)

DE MART-1 melanoma antigen immunogenic peptide M9-11.

XX

XX MART-1; M9-1; melanoma antigen recognised by T-cells; melanoma;

KW metastatic melanoma; tumour-associated antigen;

KW immunogenic peptide; diagnosis; prognosis; prophylaxis;

KW therapy; vaccine.

XX

OS Synthetic.

XX

PN W0529193-A2.

XX

PD 02-NOV-1995.

XX

PF 21-APR-1995; 95WO-US05063.

XX

PR 05-APR-1995; 95US-0417174.

PR 22-APR-1994; 94US-0231565.

XX

PA (USSH) US SEC DEPT HEALTH.

XX

PI Kawakami Y, Rosenberg SA;

XX

DR WPI; 1995-382963/49.

XX

XX DNA encoding melanoma antigens recognised by T-lymphocytes - also

PI vectors, host cells and antibodies, used to detect, treat and

PT immunise animal against melanoma.

XX

PS Example 2; Page 72; 184pp; English.

XX

CC Immunogenic peptide M9-11 is based on the melanoma antigen (MART-1)

CC (see AAR84212). M9-1 may be modified to improve immunogenicity

CC (see AAR84763-R84782) and used in medicaments (vaccines) for

CC treatment or prevention (by immunization) of melanoma.

CC Antibodies against MART-1 and its immunogenic peptides may be used

CC in the detection and isolation of MART-1 from a sample, the

CC detection of which is indicative of a disease state (melanoma

CC or metastatic melanoma).

CC See also AAR84196-R84198.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 16; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4

DB 2 VVPN 5

RESULT 15

AAU28899

ID AAU28899 standard; Peptide: 9 AA.

XX

AC AAU28899;

XX

DT 18-DEC-2001 (first entry)

XX

DE MART-1 antigen peptide M9-11.

XX

KW Human; MART-1; immunogenic; melanoma antigen recognised by T lymphocyte;

KW diagnostic; therapeutic; vaccine; melanoma; in vivo tumour recognition;

KW in vivo tumour rejection.

XX

OS Homo sapiens.

XX

PN US6270778-B1.

XX

PD 07-AUG-2001.

XX

PF 12-MAR-1999; 99US-0267439.

XX

PR 05-MAY-1998; 98US-0073138.

PR 22-APR-1994; 94US-0231565.

PR 05-APR-1995; 95US-0417174.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Kawakami Y, Rosenberg SA;

XX

DR WPI; 2001-595403/67.

XX

XX Immunogenic peptide useful in vaccines comprises specific amino acids

PI of new melanoma antigen recognised by T lymphocytes -

XX

PS Example 2; Column 37; 73pp; English.

XX

CC The invention relates to a novel immunogenic peptide comprising 5-20

CC contiguous amino acids of new melanoma antigen recognised by T

CC lymphocytes (MART-1). The peptide sequence contains at least one amino

CC acid modification of MART-1. The peptide is used in diagnostic and

CC therapeutic methods as an immunogen or vaccine to prevent or treat

CC melanoma, and for in vivo tumour recognition and rejection. AAU28888-

CC AAU29008 represent MART-1 peptide amino acid sequences, and related

CC sequences of the invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4

DB 2 VVPN 5

Search completed: September 4, 2003, 21:00:28

Job time : 21.5882 secs

CC arthritis and multiple sclerosis), muscular disorders, respiratory
 CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
 CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
 CC cardiovascular disorders (e.g. congenital heart defects, Ebstein's
 CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
 CC kidney failure and end-stage renal disease), hyperproliferative disorders
 CC (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g.
 CC septic shock, bursitis and appendicitis), allergic reactions and
 CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
 CC atherosclerosis and myocardial infarction) and cancerous diseases.
 CC Sequences ABUS4914-ABUS5699 and ABUS5748 represent human novel
 CC polypeptides of the invention.

XX Sequence 497 AA;

Query Match 90.3%; Score 28; DB 24; Length 497;

Best Local Similarity 83.3%; Pred. No. 5.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPN 6

Db 417 GSVVVPN 422

Search completed: September 4, 2003, 21:00:31
 Job time : 32.8824 secs

PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-488783/53.
 DR N-PSDB; AAS26348.
 XX
 PT New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 PS Claim 11; SEQ ID NO 1314; 980pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capability, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention.

Query Match 90.38; Score 28; DB 22; Length 497;
 Best Local Similarity 83.38; Pred. No. 5.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPP 6
 1:||||
 Db 417 GSVPN 422

RESULT 15
 ABUS5430
 ID ABUS5430 standard; Protein; 497 AA.
 XX
 AC ABUS5430;
 XX
 DT 18-MAR-2003 (first entry)
 XX
 DE Human novel polypeptide #517.
 XX
 KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 PN US2002132753-A1.
 XX
 PD 19-SEP-2002.
 XX

PF 17-JAN-2001; 2001US-0764864.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214866P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234979P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239335P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 XX (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-14744/14.
 DR N-PSDB; ABX73689.
 XX
 PT New polypeptides and nucleic acids, useful in gene therapy for
 PT treating, inhibiting or preventing e.g. neural, immune system,
 PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
 PT cardiovascular or renal disorders -
 XX
 PS Claim 11; SEQ ID NO 1314; 402pp; English.
 XX
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid

KW corneal infection; wound healing; epithelial cell proliferation;
KW skin aging; food additive; preservative; antiproliferative.
OS Homo sapiens.
PM WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198122.
PR 19-MAY-2000; 2000US-0205513.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227109.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

Query Match 90.3%; Score 28; DB 19; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
 DB 241 GAVVPN 246

RESULT 12

AAE05070
 ID AAE05070 standard; Protein: 428 AA.

AC AAE05070;

DT 10-SEP-2001 (first entry)

DE Bacillus subtilis class II EPSs protein.

KM 5-enolpyruvylshikimate-3-phosphate synthase; EPSs; glyphosate;
 KM transformed bacteria; transgenic plant; herbicide.

OS Bacillus subtilis.

PN US624876-B1.

PD 19-JUN-2001.

PF 20-AUG-1998; 98US-0137440.

PR 13-SEP-1994; 94US-0306063.

PR 07-APR-1997; 97US-0833485.

PR 31-AUG-1990; 90US-0576537.

PR 28-AUG-1991; 91US-0749611.

PA (MONS) MONSANTO CO.

PI Barry GF, Kishore GM, Padgett SR, Stallings WC;

DR WPI: 2001-407326/43.

DR N-PSDB; AAD09775.

XX DNA probe capable of use in a polymerase chain reaction for identifying

PT the presence of a target genomic DNA encoding a

PT 5-enolpyruvylshikimate-3-phosphate synthase (EPSPs) enzyme -

XX Disclosure; Fig 18; 152pp; English.

XX The present invention relates to a DNA probe capable of use in a

CC polymerase chain reaction for identifying the presence of a target

CC genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPs)

CC enzyme referred to as class II EPSPs enzyme which is tolerant to

CC glyphosate. EPSPs genes are useful in producing transformed bacteria

CC and transgenic plants which are tolerant to glyphosate herbicide. The

CC probe is useful for identifying the presence of a target genomic DNA

CC encoding a EPSPs enzyme. The present sequence is Bacillus subtilis

XX class II EPSPs protein.

XX Sequence 428 AA;

QY 1 GAVVPN 6

DB 241 GAVVPN 246

AC ABB97384;

XX 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 652.

KM Human; anti-naemic; vulnery; anti-inflammatory; immunomodulator;

KM anti-fertility; cerebroprotective; cytosolic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI: 2002-292408/33.

DR N-PSDB; ABN32570.

XX An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 20; SEQ ID NO 652; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibit e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat

CC stroke and cancer, to screen for drugs, to treat inflammatory conditions

CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention.

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

AC ABB97384;

XX 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 652.

KM Human; anti-naemic; vulnery; anti-inflammatory; immunomodulator;

KM anti-fertility; cerebroprotective; cytosolic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI: 2002-292408/33.

DR N-PSDB; ABN32570.

XX An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 20; SEQ ID NO 652; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibit e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat

CC stroke and cancer, to screen for drugs, to treat inflammatory conditions

CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention.

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

XX Sequence 469 AA;

QY 1 GAVVPN 6

DB 389 GAVVPN 394

CC EPSPS converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid
 CC to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the
 CC herbicide glyphosate. It would be useful to produce transgenic crops
 CC containing glyphosate resistance genes so that glyphosate-containing
 CC herbicides can be applied to selectively kill weeds. The novel EPSPS
 CC enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that
 CC when introduced into a plant, the plant is made glyphosate-tolerant, and
 CC EPSPS enzyme activity is not affected. These class II EPSPS enzymes are
 CC fused to a chloroplast transit peptide to target the protein into the
 CC chloroplast, which is the site for the shikimic acid pathway. In
 CC addition, the EPSPS gene is cloned into a plant under the control of a
 CC promoter such as figwort mosaic virus promoter or the cauliflower mosaic
 CC virus promoter, so that expression is enhanced.

CC
 XX
 SQ Sequence 428 AA;

Query Match 90.3%; Score 28; DB 18; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPPN 6
 ||:||||
 Db 241 GAVVPPN 246

RESULT 10
 AAW24481
 ID AAW24481 standard; Protein; 428 AA.
 XX
 AC AAW24481;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-OCT-1997 (first entry)
 XX

Class II EPSPS used for glyphosate resistant plant production.

DE
 XX 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
 KM glyphosate resistant; transgenic plant; herbicide; shikimic acid;
 KM fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
 KM tobacco.

XX
 XX Bacillus subtilis.
 XX
 XX US5633435-A.
 PN
 XX 27-MAY-1997.
 PD
 XX
 XX 13-SEP-1994; 94US-0306063.
 PF
 XX 13-SEP-1994; 94US-0306063.
 PR 31-AUG-1990; 90US-0576537.
 PR 28-AUG-1991; 91US-0749611.
 XX
 PA (MONS) MONSANTO CO.
 PI
 XX Barry GF, Kishore GM, Padgett SR, Stallings WC;
 PI
 XX WPI; 1997-297418/27.
 DR N-PSDB; AAT77317.
 DR
 XX New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
 PT used for transforming plants to produce plants which are tolerant to
 PT glyphosate herbicide
 PT
 XX
 PS Claim 7; Column 97-100; 154pp; English.
 PS
 XX AAW24481 shows the sequence of a class II 5-enolpyruvylshikimate-3-
 CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
 CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
 CC the production of herbicide resistant (glyphosate-tolerant) plants
 CC such as corn, wheat, rice, soybean, cotton, sugarbeet, oilseed
 CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
 CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.

CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 428 AA;

Query Match 90.3%; Score 28; DB 18; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPPN 6
 ||:||||
 Db 241 GAVVPPN 246

RESULT 11
 AAW71617
 ID AAW71617 standard; Protein; 428 AA.
 XX
 AC AAW71617;
 XX
 DT 14-DEC-1998 (first entry)
 DT
 XX
 XX Bacillus subtilis Class II EPSPS.
 DE
 XX 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
 KM glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
 KM resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
 KM N-phosphonomethylglycine.
 XX
 OS Bacillus subtilis.
 OS
 PN US5804425-A.
 PN
 XX 08-SEP-1998.
 PD
 XX
 XX 07-APR-1997; 97US-0833485.
 PF
 XX
 PR 13-SEP-1994; 94US-0306063.
 PR 31-AUG-1990; 90US-0576537.
 PR 28-AUG-1991; 91US-0749611.
 PR 07-APR-1997; 97US-0833485.
 XX
 XX (MONS) MONSANTO CO.
 PA
 XX Barry GF, Kishore GM, Padgett SR, Stallings WC;
 PI
 XX WPI; 1998-505657/43.
 DR N-PSDB; AAV58018.
 DR
 XX
 XX Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -
 PT useful for characterisation of the enzyme to determine inhibition
 PT data values
 PT
 XX
 PS Disclosure; Fig 18; 152pp; English.
 PS
 XX An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
 CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
 CC -G-D-R-X3- (II); -S-A-O-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
 CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
 CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
 CC produced by recombinant methods, can be used in kinetic studies to
 CC determine Ki and Km values of the enzyme for its characterisation. The
 CC enzyme is normally used for the production of 5-enolpyruvyl-3-
 CC phosphoshikimic acid in plants, and most forms of the enzyme are
 CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
 CC Inhibition data enables more accurate values of concentrations of
 CC herbicide to be used when growing the plant without being detrimental to
 CC it. This enables the plant to be grown in the presence of the herbicide,
 CC being used to inhibit the growth of undesired plants. The present
 CC sequence represents a Class II EPSPS from bacterial isolate
 CC Bacillus subtilis.
 XX
 SQ Sequence 428 AA;

CC		at ftp.wipo.int/pub/published_pct_sequences.
XX		
SQ	Sequence	337 AA;
	Query Match	90.3%; Score 28; DB 22; Length 337;
	Best Local Similarity	83.3%; Pred. No. 3.7e+02;
	Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY	1 GAVVPN 6	
	:	
DB	221 GSVPN 226	
	RESULT 9	
AAMW34686	ID	AAMW34686 standard; Protein: 428 AA.
XX	AC	AAMW34686;
XX	DT	17-FEB-1998 (first entry)
DE	Class II EPS synthase (EPSs) from Bacillus subtilis.	
XX		
KM	5-enolpyruvylshikimate-3-phosphate synthase; EPSs; glyphosate;	
KW	shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;	
KV	5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;	
KW	glyphosate resistance gene; glyphosate-tolerance; promoter.	
XX	Bacillus subtilis.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	Region	/190..194
FT		/label=characteristic_region
FT		/note="see AAW34690"
FT	Region	/17..20
FT		/label=characteristic_region
FT		/note="see AAW34691"
FT	Region	/164..168
FT		/label=characteristic_region
FT		/note="see AAW34692"
FT	Region	/257..260
FT		/label=characteristic_region
FT		/note="see AAW34693"
PX	US5627061-A.	
XX		
PD	06-MAY-1997.	
XX		
PF	07-JUN-1995;	95US-0476008.
XX		
PR	13-SEP-1994;	94US-0306063.
PR	31-AUG-1990;	90US-0576537.
PR	28-AUG-1991;	91US-0749611.
XX	(MONS) MONSANTO CO.	
PA		
PI	Barry GF, Kishore GM, Padgett SR, Stallings WC;	
XX		
XX	WPI; 1997-271315/24.	
DR	N-PDB; AAT93793.	
XX		
PT	Production of glyphosate-herbicide tolerant plants - using DNA	
PT	encoding class II 5-enolpyruvylshikimate-3-phosphate synthase	
PT	enzyme(s)	
PS	Claim 6; Fig 18; 151pp; English.	
XX		
CC	AAW34683-89 represent a new class of glyphosate-tolerant	
CC	5-enolpyruvylshikimate-3-phosphate synthases (EPSs). These novel	
CC	EPS enzymes have little homology with known Class I EPS enzymes, and	
CC	belong to a new class, Class II. The present sequence was isolated from	
CC	Bacillus subtilis. The EPS enzymes are part of the shikimic	
CC	acid pathway, which leads to the biosynthesis of aromatic compounds	

PS Example 1; SEQ ID No 19445; 1069pp; English.
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SARPO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 79 AA:

Query Match 96.8%; Score 30; DB 22; Length 79;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
|||:|
DB 40 GAVIPN 45

RESULT 6
AAV34508
ID AAV34508 standard; Protein: 867 AA.

XX AAV34508;

DT 20-MAR-2003 (updated)
DT 25-AUG-1999 (first entry)

DE Porphyromonas gingivalis protein PG47.

XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic.

XX Porphyromonas gingivalis.

XX W09929870-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-AU01023.

XX 10-DEC-1997; 97AU-0000839.

XX 31-DEC-1997; 97AU-0001182.

XX 30-JAN-1998; 98AU-0001546.

XX 10-MAR-1998; 98AU-0002264.

XX 09-APR-1998; 98AU-0002911.

XX 23-APR-1998; 98AU-0003128.

XX 05-MAY-1998; 98AU-0003338.

XX 22-MAY-1998; 98AU-0003654.

XX 29-JUL-1998; 98AU-0004917.

XX 30-JUL-1998; 98AU-0004963.

XX 04-AUG-1998; 98AU-0005028.

XX (CSLC-) CSL LTD.

XX Agius CT, Barr IG, Hocking DM, Margelits MB, Patterson MA;
PI Ross BC, Rochel LJ, Webb EA;
XX

DR WPI: 1999-385613/32.
DR N-PSDB; AAX91726.
XX
PT Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis
XX
PS Claim 1; Page 491-492; 588pp; English.

XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAV34318 to
CC AAX34583. AAX91802 to AAX91969 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.
XX (Updated on 20-MAR-2003 to correct PR field.)
XX

SQ Sequence 867 AA;

Query Match 93.5%; Score 29; DB 20; Length 867;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
|||:|
DB 619 GAVIPN 624

RESULT 7
AAV34381
ID AAV34381 standard; Protein: 875 AA.

XX AAV34381;

DT 20-MAR-2003 (updated)
DT 25-AUG-1999 (first entry)

DE Porphyromonas gingivalis protein PG47.

XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic.

XX Porphyromonas gingivalis.

XX W09929870-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-AU01023.

XX 10-DEC-1997; 97AU-0000839.

XX 31-DEC-1997; 97AU-0001182.

XX 30-JAN-1998; 98AU-0001546.

XX 10-MAR-1998; 98AU-0002264.

XX 09-APR-1998; 98AU-0002911.

XX 23-APR-1998; 98AU-0003128.

XX 05-MAY-1998; 98AU-0003338.

XX 22-MAY-1998; 98AU-0003654.

XX 29-JUL-1998; 98AU-0004917.

XX 30-JUL-1998; 98AU-0004963.

XX 04-AUG-1998; 98AU-0005028.

XX (CSLC-) CSL LTD.

XX Agius CT, Barr IG, Hocking DM, Margelits MB, Patterson MA;
PI Ross BC, Rochel LJ, Webb EA;
XX WPI: 1999-385613/32.
DR N-PSDB; AAX91599.
XX
PT Antigenic Porphyromonas gingivalis peptides for preventing

XX Claim 24; Page 21; 53pp; English.

XX The present invention relates to a composition for the treatment of

CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or

CC their biological equivalents. The composition is useful not only for

CC improving the cosmetic appearance of skin, but for the treatment of

CC various skin disorders associated with reduced elastin formation and for

CC numerous other tissue disorders especially of blood vessels including

CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary

CC thrombosis, chronic obstructive pulmonary disease and restenosis post-

CC angioplasty. They may also be useful in coating surgical appliances i.e.

CC stents. The use of peptides and peptide fragments which mimic those found

CC in elastin overcomes the problem of solubility of the dense cross-linked

CC structure of elastin itself, which renders it difficult to use

CC therapeutically. The peptides are found to improve elastin production in

CC treated tissues. The present sequence is ligamentum nuchae hydrolysed

CC elastin peptide (HEP), VVPN derived peptide.

XX

SQ Sequence 8 AA:

Query Match 100.0%; Score 31; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6

Db 2 GAVVPN 7

RESULT 4

AAE16730

ID AAE16730 standard; peptide; 8 AA.

XX

AC AAE16730;

XX

DT 09-APR-2002 (first entry)

XX

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.

XX

XX Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;

KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;

KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;

KW hypertensive; chronic obstructive pulmonary disease; antiatherosclerotic;

KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;

KW cyclic.

XX

XX Ligamentum nuchae.

OS Synthetic.

OS

XX

XX

FT Key Location/Qualifiers

FT Modified-site 1

FT /note- "This residue forms a cyclic structure with Cys

FT at position 8 with copper as a chelating agent"

FT Modified-site 8

FT /note- "This residue forms a cyclic structure with Cys

FT at position 1 with copper as a chelating agent"

FT

XX WO200191700-A2.

XX

XX

PD 06-DEC-2001.

XX

XX

PF 30-MAY-2001; 2001WO-US17384.

XX

XX

PR 30-MAY-2000; 2000US-0580110.

XX

PR 30-MAY-2000; 2000US-0580156.

XX

PR 30-MAY-2000; 2000US-0580893.

XX

PR 30-MAY-2000; 2000US-0584001.

XX

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.

PA

XX

PI Mlts TF, Sandberg LB, Jimenez F;

XX

DR WPI; 2002-106259/14.

XX

PT Composition for improving the elasticity of tissue comprises peptide

XX substances which mimic the sequences found in elastin

XX

PS Claim 24; Page 21; 53pp; English.

XX

XX The present invention relates to a composition for the treatment of

CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or

CC their biological equivalents. The composition is useful not only for

CC improving the cosmetic appearance of skin, but for the treatment of

CC various skin disorders associated with reduced elastin formation and for

CC numerous other tissue disorders especially of blood vessels including

CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary

CC thrombosis, chronic obstructive pulmonary disease and restenosis post-

CC angioplasty. They may also be useful in coating surgical appliances i.e.

CC stents. The use of peptides and peptide fragments which mimic those found

CC in elastin overcomes the problem of solubility of the dense cross-linked

CC structure of elastin itself, which renders it difficult to use

CC therapeutically. The peptides are found to improve elastin production in

CC treated tissues. The present sequence is ligamentum nuchae hydrolysed

CC elastin peptide (HEP), VVPN derived cyclic peptide.

XX

SQ Sequence 8 AA:

Query Match 100.0%; Score 31; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6

Db 2 GAVVPN 7

RESULT 5

AAU58250

ID AAU58250 standard; protein; 79 AA.

XX

AC AAU58250;

XX

DT 27-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #19146.

XX

XX SAPHO syndrome; synovitis; acne; pustulosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

XX

XX Propionibacterium acnes.

OS

OS

XX

XX

PN WO200181581-A2.

XX

XX

PD 01-NOV-2001.

XX

XX

PF 20-APR-2001; 2001WO-US12865.

XX

XX

PR 21-APR-2000; 2000US-199047P.

XX

PR 02-JUN-2000; 2000US-208841P.

XX

PR 07-JUL-2000; 2000US-216747P.

XX

XX

XX (CORI-) CORIXA CORP.

XX

XX

PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX

DR WPI; 2001-616774/71.

XX

DR N-PSDB; AAS59589.

XX

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

XX

PI Mlts TF, Sandberg LB, Jimenez F;
 XX WPI: 2002-106259/14.
 DR
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 PS
 XX Claim 24; Page 21; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 CC
 XX Sequence 6 AA:
 SO
 Query Match 100.0%; Score 31; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GAVVPN 6
 |||||
 Db 1 GAVVPN 6
 |||||
 |||||

RESULT 2
 AAEI6724
 ID AAEI6724 standard; peptide: 6 AA.
 XX
 AC AAEI6724;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #4.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 6 /note="C-terminal amide"
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGINEERING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;

XX
 DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 PS
 XX Claim 24; Page 21; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 CC
 XX Sequence 6 AA:
 SO
 Query Match 100.0%; Score 31; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GAVVPN 6
 |||||
 Db 1 GAVVPN 6
 |||||
 |||||

RESULT 3
 AAEI6727
 ID AAEI6727 standard; peptide: 8 AA.
 XX
 AC AAEI6727;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 6 /note="C-terminal amide"
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGINEERING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 PT WPI: 2002-106259/14.
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 30.8824 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893D-57

Perfect score: 31
Sequence: 1 GAVPN 6

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	6	23	AAE16722
2	31	100.0	6	23	AAE16724
3	31	100.0	8	23	AAE16727
4	31	100.0	8	23	AAE16730
5	30	96.8	79	22	AAU58250
6	29	93.5	867	20	AAI34508
7	29	93.5	875	20	AAI34381
8	28	90.3	337	22	ABBS9302
9	28	90.3	428	18	AAW34686

10	28	90.3	428	18	AAW24481
11	28	90.3	428	19	AAW71617
12	28	90.3	428	22	AAE05070
13	28	90.3	469	23	ABB97384
14	28	90.3	497	22	AAU16361
15	28	90.3	497	24	ABU55430
16	28	90.3	534	22	AAU15908
17	28	90.3	534	24	ABU54977
18	28	90.3	547	23	ABBS9206
19	28	90.3	707	22	AAW93718
20	28	90.3	707	22	AAW93234
21	28	90.3	771	20	AAI34574
22	28	90.3	786	20	AAI34431
23	28	90.3	893	23	ABJ19186
24	28	90.3	901	23	ABJ40544
25	28	90.3	944	22	AAU03569
26	28	90.3	1938	24	ABB98398
27	27	87.1	75	24	ABP75905
28	27	87.1	89	22	AAU59414
29	27	87.1	101	22	ABR06357
30	27	87.1	102	22	AAU45639
31	27	87.1	103	22	ABR96076
32	27	87.1	103	22	AAW95385
33	27	87.1	208	23	ABP25631
34	27	87.1	231	21	AAI46950
35	27	87.1	247	22	AAI89920
36	27	87.1	252	20	AAI37064
37	27	87.1	271	24	ABP77868
38	27	87.1	359	23	ABBS3778
39	27	87.1	375	22	ABG26783
40	27	87.1	394	22	ABG26781
41	27	87.1	395	21	AAI46949
42	27	87.1	441	23	ABH76130
43	27	87.1	637	21	AAI26944
44	27	87.1	655	23	ABP40686
45	27	87.1	663	22	ABBS9368

ALIGNMENTS

RESULT 1	AAE16722	standard; peptide; 6 AA.
ID	AAE16722	
XX	AAE16722;	
AC	09-APR-2002	(first entry)
XX		
DT		
XX		
DE		Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #2.
XX		
KW		Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW		vasotrophic; hypertension; skin disorder; dermatological;
KW		coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW		hypotensive; chronic obstructive pulmonary disease; antihypertensive;
KW		restenosis post-angioplasty; antilipidemic; antithrombotic; thrombolytic.
XX		
OS		Ligamentum nuchae.
XX		
OS		Synthetic.
XX		
PN		W0200191700-A2.
XX		
PD		06-DEC-2001.
XX		
PF		30-MAY-2001; 2001WO-US17384.
XX		
PR		30-MAY-2000; 2000US-0580110.
XX		
PR		30-MAY-2000; 2000US-0580156.
XX		
PR		30-MAY-2000; 2000US-0580893.
XX		
PR		30-MAY-2000; 2000US-0584001.
XX		
PA		(CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX		

Class II EPSPs use
Bacillus subtilis
Bacillus subtilis
Novel human protei
Human novel secret
Human novel polype
Human novel secret
Human novel polype
Herbicideally activ
Human polypeptide,
Human protein sequ
Porphyromonas ging
Porphyromonas ging
pathogen specific
Staphylococcus epi
P. patens active p
Streptomyces virid
Human secretory po
Propionibacterium
Human zinc finger
Propionibacterium
Human testicular a
Human reproductive
Streptococcus poly
Arabidopsis thalia
C glutamicum prote
Protein involved i
N. gonorrhoeae and
Lactococcus lactis
Novel human diagno
Novel human diagno
Arabidopsis thalia
zygococcus hellenic
Soybean auxin tran
Staphylococcus epi
Drosophila melanog

RESULT 13

072702

ID 072702

PRELIMINARY;

PRT;

200 AA.

AC 072702;

01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE NSS protein (Fragment).

GN NSS.

OS Bunyavirus sp. 'strain SNS'.

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.

OX NCBI_TaxID=75386;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SNS;

RA MEDLINE=98033195; PubMed=9367372;

RA Sali A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,

RA Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus isolates."

RT Isolates."

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-SNS;

RA Prehaud C.J.N.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: Y12755; CAA73289.1; -.

FT NON_TER 1 1.

FT NON_TER 200 200

SQ SEQUENCE 200 AA; 22472 MW; E7551856FE61CD88 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5

Db 187 AAVPN 191

RESULT 14

073542

ID 073542

PRELIMINARY;

PRT;

200 AA.

AC 073542;

DT 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE NSS protein (Fragment).

GN NSS.

OS Rift valley fever virus (RVFV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

OX NCBI_TaxID=11588;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar Mg990;

RX MEDLINE=98033195; PubMed=9367372;

RA Sali A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,

RA Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus isolates."

RT Isolates."

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar Mg990;

RA Prehaud C.J.N.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: Y12746; CAA73280.1; -.

FT NON_TER 1 1.

FT NON_TER 200 200

SQ SEQUENCE 200 AA; 22514 MW; FBA810C78CB64A20 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5

Db 187 AAVPN 191

RESULT 15

073539

ID 073539

PRELIMINARY;

PRT;

200 AA.

AC 073539;

DT 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE NSS protein (Fragment).

GN NSS.

OS Rift valley fever virus (RVFV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

OX NCBI_TaxID=11588;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar D38457;

RX MEDLINE=98033195; PubMed=9367372;

RA Sali A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,

RA Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus isolates."

RT Isolates."

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar D38457;

RA Prehaud C.J.N.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: Y12743; CAA73277.1; -.

FT NON_TER 1 1.

FT NON_TER 200 200

SQ SEQUENCE 200 AA; 22276 MW; 34A2CA0161EBB4DB CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5

Db 187 AAVPN 191

Search completed: September 4, 2003, 21:06:37
Job time : 22.6056 secs

RP SEQUENCE FROM N.A.
RC STRAIN-B Eg93;
RX MEDLINE-9803195; PubMed-9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
Thiongane Y., Bouloy M.;
RT "Variability of the NSS protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN-B Eg93;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12747; CAA73281.1; -.
FT NON_TER 1 1
FT SEQUENCE 200 AA; 22592 MW; C7C76FFEBB3C98DF CRC64;
SQ
Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPN 5
DB 187 AVPN 191
RESULT 10
O73548 PRELIMINARY; PRT: 200 AA.
ID 073548:
AC 073548:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MP12;
RX MEDLINE-9803195; PubMed-9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
Thiongane Y., Bouloy M.;
RT "Variability of the NSS protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MP12;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12754; CAA73288.1; -.
FT NON_TER 1 1
FT SEQUENCE 200 AA; 22564 MW; D8C87FFEBB2B8DFD CRC64;
SQ
Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPN 5
DB 187 AVPN 191
RESULT 11
O73537 PRELIMINARY; PRT: 200 AA.
ID 073537:
AC 073537:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update).

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar B1976;
RX MEDLINE-9803195; PubMed-9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
Thiongane Y., Bouloy M.;
RT "Variability of the NSS protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar B1976;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12741; CAA73275.1; -.
FT NON_TER 1 1
FT SEQUENCE 200 AA; 22533 MW; 271D6D130A4F7E9B CRC64;
SQ
Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPN 5
DB 187 AVPN 191
RESULT 12
O73547 PRELIMINARY; PRT: 200 AA.
ID 073547:
AC 073547:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H Eg93A;
RX MEDLINE-9803195; PubMed-9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
Thiongane Y., Bouloy M.;
RT "Variability of the NSS protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H Eg93A;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12751; CAA73285.1; -.
FT NON_TER 1 1
FT SEQUENCE 200 AA; 22654 MW; 823F356F756E3AF3 CRC64;
SQ
Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPN 5
DB 187 AVPN 191

```

RP SEQUENCE FROM N.A.
RC STRAIN-H DA7311;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12748; CAA73282.1; -.
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22572 MW; F3F3740A939F3C68 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVFN 5
Db 187 AVVFN 191

RESULT 6
073540
ID 073540 PRELIMINARY; PRT; 200 AA.
AC 073540;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSs protein (Fragment).
GN NSs.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ar D38661;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
RT isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ar D38661;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12744; CAA73278.1; -.
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22544 MW; 485876F739FDA45F CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVFN 5
Db 187 AVVFN 191

RESULT 7
08JUF9
ID 08JUF9 PRELIMINARY; PRT; 200 AA.
AC 08JUF9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE NSs protein (Fragment).
GN NSs.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA01-1322;

```

```

RA Miller B.R., Godsey M.S., Crabtree M.B., Al-Mazrao Y., Al-Jeffri M.H.,
RA Adoon A.M., Al-Seghayer S.M., Ksiazek T.G.;
RT "Out of Africa: Isolation and Genetic Characterization of Rift Valley
RT Fever Virus from Redes arabiensis from the Asir Region of the Kingdom
RT of Saudi Arabia.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF393744; AAM73694.1; -.
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22530 MW; 217BE1B99CB4C6B6 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVFN 5
Db 187 AVVFN 191

RESULT 8
073541
ID 073541 PRELIMINARY; PRT; 200 AA.
AC 073541;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSs protein (Fragment).
GN NSs.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ar M6811;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
RT isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ar M6811;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12745; CAA73279.1; -.
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22502 MW; EA7EFB79DB833419 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVFN 5
Db 187 AVVFN 191

RESULT 9
073543
ID 073543 PRELIMINARY; PRT; 200 AA.
AC 073543;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSs protein (Fragment).
GN NSs.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]

```

OY 1 AVPN 5
 DB 42 AVPN 46

RESULT 2

O92X31 PRELIMINARY; PRT; 161 AA.
 ID O92X31:
 AC O92X31:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative oxidoreductase protein.
 GN RB0131 OR SM20131.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Flann T.M., Weldner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhoeft F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603642; CAC48531.1; -
 DR InterPro; IPR002888; 2Fe-2S_bind.
 DR InterPro; IPR006058; 2Fe2S_ferredoxin.
 DR Pfam; PF00111; fer2.1.
 DR Pfam; PF01799; fer2.2; 1.
 DR ProDom; PD186071; 2Fe-2S_bind.1.
 DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 1.
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 161 AA; 17182 MW; 7128EAFCFBFC1FF7 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 161;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 17 AVPN 21

RESULT 3

O9ZG06 PRELIMINARY; PRT; 174 AA.
 ID O9ZG06:
 AC O9ZG06:
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Type-I signal peptidase Sipa.
 GN Sipa.
 OS Staphylococcus carnosus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1281;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TM300;
 RA Matzen A., Frendl R.;
 RT "Cloning and characterization of the type-I signal peptidase of
 RT Staphylococcus carnosus";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF089862; AAD09010.1; -
 DR HSSP; P00803; 1812.
 DR InterPro; IPR000508; SigPase.
 DR Pfam; PF00461; Peptidase_S26; 1.

SQ SEQUENCE 174 AA; 19931 MW; 965B7A8FDB818C2B CRC64;

Query Match 100.0%; Score 25; DB 2; Length 174;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 29 AVPN 33

RESULT 4

O73538 PRELIMINARY; PRT; 200 AA.
 ID O73538:
 AC O73538:
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NSS protein (Fragment).
 GN NSS.
 OS Rift valley fever virus (RVFV).
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 OX NCBI_TaxID=11588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Ar D104769;
 RX MEDLINE=98033195; PubMed=9367372;
 RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
 RA Thiongane Y., Bouloy M.;
 RT "Variability of the NSS protein among rift valley fever virus
 RT isolates";
 RL J. Gen. Virol. 78:2853-2858(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Ar D104769;
 RA Prehaud C.J.N.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y12742; CAA73276.1; -
 FT NON_TER 1 1
 FT NON_TER 200 200
 SQ SEQUENCE 200 AA; 22558 MW; 1DE3740A039F3C6D CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 187 AVPN 191

RESULT 5

O73544 PRELIMINARY; PRT; 200 AA.
 ID O73544:
 AC O73544:
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NSS protein (Fragment).
 GN NSS.
 OS Rift valley fever virus (RVFV).
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 OX NCBI_TaxID=11588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H DA7311;
 RX MEDLINE=98033195; PubMed=9367372;
 RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
 RA Thiongane Y., Bouloy M.;
 RT "Variability of the NSS protein among rift valley fever virus
 RT isolates";
 RL J. Gen. Virol. 78:2853-2858(1997).
 RN [2]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 20.5147 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893d-56
Perfect score: 25
Sequence: 1 AVVPN 5

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	134	2	Q53805 streptomyc
2	25	100.0	161	16	Q92X31 Rhizobium m
3	25	100.0	174	2	Q9ZG06 staphylococ
4	25	100.0	200	12	Q73538 rftt valley
5	25	100.0	200	12	Q73544 rftt valley
6	25	100.0	200	12	Q73540 rftt valley
7	25	100.0	200	12	Q8JUP9 rftt valley
8	25	100.0	200	12	Q73541 rftt valley
9	25	100.0	200	12	Q73543 rftt valley
10	25	100.0	200	12	Q73548 rftt valley
11	25	100.0	200	12	Q73537 rftt valley
12	25	100.0	200	12	Q73547 rftt valley
13	25	100.0	200	12	Q72702 bunyavirus
14	25	100.0	200	12	Q73542 rftt valley
15	25	100.0	200	12	Q73539 rftt valley
16	25	100.0	200	12	Q72701 lunyo virus

17	25	100.0	200	12	Q73545 rftt valley
18	25	100.0	200	12	Q73549 rftt valley
19	25	100.0	200	12	Q73535 rftt valley
20	25	100.0	200	12	Q73536 rftt valley
21	25	100.0	200	12	Q73546 rftt valley
22	25	100.0	217	10	Q9ZK25 populus tre
23	25	100.0	217	17	Q8TY62 methanopyru
24	25	100.0	229	10	Q9UZK5 pyrococcus
25	25	100.0	234	10	Q8S7U8 oryza sativ
26	25	100.0	242	4	Q9HIG7 homo sapien
27	25	100.0	247	5	Q8IR10 dtrosophila
28	25	100.0	284	11	Q9LWL7 mus musculu
29	25	100.0	292	16	Q9Z9E2 listeria in
30	25	100.0	292	16	Q8Y543 listeria mo
31	25	100.0	296	2	Q86189 erwinia chr
32	25	100.0	301	2	Q88169 enterococu
33	25	100.0	312	10	Q9LJ09 oryza sativ
34	25	100.0	326	10	Q9SGK8 lycopersico
35	25	100.0	332	11	Q8K0U1 mus musculu
36	25	100.0	334	16	Q8YS25 anabaena sp
37	25	100.0	334	16	Q9CPV7 lactococcus
38	25	100.0	348	16	Q8C2X4 yeisnia pe
39	25	100.0	351	16	Q9PF30 xyella fas
40	25	100.0	351	16	Q8P125 xanthomonas
41	25	100.0	353	16	Q8P6S6 xanthomonas
42	25	100.0	363	17	Q9YCS0 aeropyrum p
43	25	100.0	365	10	Q8VZQ1 arabidopsis
44	25	100.0	381	5	Q9GT27 brugia mala
45	25	100.0	386	12	Q9LWNA hydrangea m

ALIGNMENTS

RESULT 1
ID Q53805 PRELIMINARY; PRT; 134 AA.
Q53805
AC Q53805;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Pre-phenomycin precursor.
GN PHM.
OS Streptomyces roseoverficillatus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales.
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=66429;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC29809;
RA MEDLINE=92104838; PubMed=1761418;
RA Muramatu R., Abe S., Hayashi H., Yamaguchi K., Jinda K., Sakano K.,
RA Inouye Y., Nakamura S.;
RT "Complete amino acid sequence of phenomycin, an antitumor polypeptide
RT antibiotic.";
RT J. Antibiot. 44:1222-1227(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC29809;
RA MEDLINE=94230131; PubMed=8175490;
RA Sakata N., Oka T., Ikono S., Horii M.;
RT "Nucleotide sequence of the phenomycin gene from Streptovorticillium
RT baldacci M564-cl.";
RL J. Antibiot. 47:370-371(1994).
DR EMBL; D17759; BAA04606.1; .
KW SIGNAL.
FT SIGNAL.
FT CHAIN 46 134 POTENTIAL.
SQ SEQUENCE 134 AA; 13751 MW; 577ADD7CFE503418E CRC64;
Query Match 100.0%; Score 25; DB 2; Length 134;
Best local Similarity 100.0%; Pred. No. 2; le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR002197; HTH_Fis.
 DR InterPro: IPR002078; Sig54_Interact.
 DR Pfam: PF01590; GAF.1.
 DR Pfam: PF02954; HTH_8; 1.
 DR Pfam: PF00158; Sigma54_activat; 1.
 DR PRINTS: PR01590; HTHFIS.
 DR SMART: SM00382; AAA; 1.
 DR SMART: SM00065; GAF; 1.
 DR TIGRfams: TIGR01199; HTH_fis; 1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE: PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE: PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE: PS50045; SIGMA54_INTERACT_4; 1.
 KW Nitrogen fixation; Transcription regulation; Activator;
 KW ATP-binding; DNA-binding.
 FT DOMAIN 205 433 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 FT DOMAIN 434 582 INTER-DOMAIN LINKER.
 FT DOMAIN 583 625 C-TERMINAL DNA-BINDING DOMAIN.
 FT NP_BIND 233 240 ATP (POTENTIAL).
 FT NP_BIND 296 305 ATP (POTENTIAL).
 FT METAL 447 447 BY SIMILARITY.
 FT METAL 452 452 BY SIMILARITY.
 FT DNA_BIND 597 616 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 625 AA; 67855 MW; D060EAB6CBIFFEC3 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 625;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPN 5
 DB 96 AVPN 100

RESULT 15
 FLEY_CAUCR STANDARD; PRT; 954 AA.
 AC P15345; P15346;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Regulatory protein flaeY.
 GN flaeY OR flaeY/FLAY OR CCI1465.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=153892;
 RN [1]
 RP PRELIMINARY SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=89178645; PubMed=2648000;
 RA Kaplan J.B., Dingwall A., Bryan R., Chamber R., Shapiro L.,
 RT "Temporal regulation and overlap organization of two Caulobacter
 RT flagellar genes.";
 RL J. Mol. Biol. 205:71-83(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocky I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -!- FUNCTION: FUNCTIONS IN TRANS TO MODULATE THE LEVEL OF
 CC TRANSCRIPTION OF THE FLAGELLIN GENES AND SEVERAL GENES ENCODING
 CC CHEMOTAXIS FUNCTIONS. IT IS ITSELF TEMPORALLY CONTROLLED.

CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS. IT ENCODES FOR TWO SEPARATE ADJACENT ORFS, flae AND
 CC FLAY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: AE005820; AKR23445.1; -
 CC PIR: A87431; A87431.
 DR TIGR: CCI1465; -
 KW Flagella; Transcription regulation; Trans-acting factor;
 KW Complete proteome.
 SQ SEQUENCE 954 AA; 97979 MW; BC56554EBFB8605D CRC64;

Query Match 100.0%; Score 25; DB 1; Length 954;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPN 5
 DB 36 AVPN 40

Search completed: September 4, 2003, 21:01:41
 Job time : 5.33824 secs

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE016804; AA010633.1; -
 DR HAMAP; MF_00534; -; 1.
 DR InterPro; IPR004522; AsnS.
 DR InterPro; IPR004364; tRNA-synt_2.
 DR InterPro; IPR002312; tRNA-synt_2.
 DR InterPro; IPR004365; tRNA-anti.
 DR InterPro; IPR006195; tRNA_ligase_II.
 DR Pfam; PF00152; tRNA-synt_2; 1.
 DR Pfam; PF01336; tRNA-anti; 1.
 DR PRINTS; PR01042; TRNASYTHASP.
 DR TIGRFS; TIGR00457; asns; 1.
 DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 466 AA; 52406 MW; AA577F799612F58B CRC64;

Query Match 100.0%; Score 25; DB 1; Length 466;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 |||||
 DB 52 AVPN 56

RESULT 13
 NIFA_AZOI STANDARD; PRT; 624 AA.
 AC P54929;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nit-specific regulatory protein.
 GN NIFA.
 OS Azospirillum lipoferum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Azospirillum.
 OX NCBI_TaxID=193;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shigematsu T., Hidaka M., Masaki H., Uozumi T.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NIFA, A TRANSCRIPTIONAL ACTIVATOR, IS REQUIRED FOR
 CC ACTIVATION OF MOST NIF OPERONS, WHICH ARE DIRECTLY INVOLVED IN
 CC NITROGEN FIXATION. NIFA INTERACTS WITH SIGMA-54.
 CC -1- SIMILARITY: Contains 1 sigma-54 factor interaction ATP-binding
 CC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D13799; BAA02956.1; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003018; GAF.
 DR InterPro; IPR002197; HTH_Fis.
 DR InterPro; IPR002078; Sig54_interact.
 DR Pfam; PF01590; GAF; 1.
 DR Pfam; PF02954; HTH_8; 1.
 DR Pfam; PF00158; Sigma54_activat; 1.
 DR PRINTS; PR01590; HTHFIS.

DR SMART; SM00382; AAA; 1.
 DR SMART; SM00065; GAF; 1.
 DR TIGRFS; TIGR01199; HTH_fis; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE; PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE; PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
 KW Nitrogen fixation; Transcription regulation; Activator;
 KW ATP-binding; DNA-binding; Metal-binding.
 KW DOMAIN 203 431 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 FT DOMAIN 432 581 INTER-DOMAIN LINKER.
 FT DOMAIN 582 624 C-TERMINAL DNA-BINDING DOMAIN.
 FT NP_BIND 231 238 ATP (POTENTIAL).
 FT NP_BIND 294 303 ATP (POTENTIAL).
 FT METAL 445 445 BY SIMILARITY.
 FT METAL 450 450 BY SIMILARITY.
 FT METAL 456 450 BY SIMILARITY.
 FT DNA_BIND 596 615 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 624 AA; 67119 MW; 223A36AD44E7016 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 624;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 |||||
 DB 96 AVPN 100

RESULT 14
 NIFA_AZOBR STANDARD; PRT; 625 AA.
 ID NIFA_AZOBR
 AC P30667;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nit-specific regulatory protein.
 GN NIFA.
 OS Azospirillum brasilense.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Azospirillum.
 OX NCBI_TaxID=192;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sp7;
 RX MEDLINE=92140038; PubMed=1779763;
 RA Liang Y.Y., Kaminski P.A., Elmerich C.;
 RT Identification of a nifA-like regulatory gene of Azospirillum
 RT brasilense Sp7 expressed under conditions of nitrogen fixation and in
 RT the presence of air and ammonia.
 RL Mol. Microbiol. 5:2735-2744(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=93122522; PubMed=1362170;
 RA Liang Y.Y., de Zamaroczy M., Arsene F., Paquelin A., Elmerich C.;
 RT Regulation of nitrogen fixation in Azospirillum brasilense Sp7;
 RT Involvement of nifA, glnA and glnB gene products.
 RL FEBS Microbiol. Lett. 79:113-119(1992).
 CC -1- FUNCTION: NIFA, A TRANSCRIPTIONAL ACTIVATOR, IS REQUIRED FOR
 CC ACTIVATION OF MOST NIF OPERONS, WHICH ARE DIRECTLY INVOLVED IN
 CC NITROGEN FIXATION. NIFA INTERACTS WITH SIGMA-54.
 CC -1- SIMILARITY: Contains 1 sigma-54 factor interaction ATP-binding
 CC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60714; CAA43126.1; -
 DR PIR; S18420; S18420.

```

AC Q9KSF9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Asparaginyl-tRNA synthetase (EC 6.1.1.22) (Asparagine--tRNA ligase)
DE (AsnRS).
GN ASNS OR VC1297.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Esmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "RNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Mature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004209; AAF94456.1; ALT_INIT.
DR HSSP: Q52428; 1B8A.
DR TIGR: VC1297; -.
DR HAMAP: MF_00534; -. 1.
DR InterPro: IPR004522; AsnS.
DR InterPro: IPR004364; tRNA-synt_2.
DR InterPro: IPR002312; tRNA-synt_asp.
DR InterPro: IPR006195; tRNA_ligase_II.
DR Pfam: PF00152; tRNA-synt_2; 1.
DR Pfam: PF01336; tRNA_ant1; 1.
DR PRINTS: PRO1042; TRNASTNTASP.
DR TIGRFAMs: TIGR00457; asns; 1.
DR PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 466 AA; 52455 MW; BD7E47B85417BD2 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 52 AVVPN 56

```

```

DE (AsnRS).
GN ASNS OR VP1893.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 221063 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.,
RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancel: 361:743-749(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP005079; BAC60156.1; -.
DR HAMAP: MF_00534; -. 1.
DR PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 466 AA; 52590 MW; 2B2F9B2D75E6C56 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 52 AVVPN 56

RESULT 12
SYN_VIBVU STANDARD: PRT; 466 AA.
AC Q8DAF0;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Asparaginyl-tRNA synthetase (EC 6.1.1.22) (Asparagine--tRNA ligase)
DE (AsnRS).
GN ASNS OR VV12254.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----

```


DE Dihydrodipicolinate synthase 1, chloroplast precursor (EC 4.2.1.52)
 DE (DHDPS 1).
 GN DHDPS1 OR DHDPS OR DHDPS1 OR AT3G60880 OR T4C21_290.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_taxid=3702;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94325478; PubMed=8049377;
 RA Vauterin M., Jacobs M.,
 RT *Isolation of a poplar and an Arabidopsis thaliana dihydrodipicolinate
 RT synthase cDNA clone.";
 RL Plant Mol. Biol. 25:545-550(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unselid M.,
 RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,
 RA Deleney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choiane N., Artiguenave F., Robert C., Brothier P.,
 RA Wincker P., Catolico L., Weissbach J., Sautin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbech E., Dizonex H., Erle H., Jordan N., Brandt P., Nakamura G.,
 RA Weidemann R., Kranz H., Voss H., Holland R., Brandt P., Simionati B.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Nordstiek G.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Scharte M., Schoen O., Bauges M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottensmeyer B., Duchemin D.,
 RA Cooke R., Lande M., Berger-Liauro C., Fumelle B., Masny D.,
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Morfort A., Argitlon A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Jenkins J.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T.H., Rizzo M., Wals A., Uteback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Kiyokawa C., Kohara T., Idesawa K., Kawashima K., Kishida Y.,
 RA Nakayama S., Nakazaki N., Shimo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
 RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RA thaliana.";
 RL Nature 408:820-822(2000).
 RN [3]
 RP SEQUENCE OF 1-158 FROM N.A.
 RA Vauterin M.,
 RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: L-aspartate 4-semialdehyde + pyruvate =
 CC dihydrodipicolinate + 2 H(2)O.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 CC semialdehyde; first step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the DHDPS family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X72971, CAB45642.1, -
 DR EMBL: AL162295, CAB2692.1, -
 DR EMBL: X98080, CAB6703.1, -
 DR PIR: S46304, S46304.
 DR PIR: T47899, T47899.
 DR HSSP: P05640, IDHP.

DR InterPro: IPR005263; DAPA.
 DR InterPro: IPR002220; DHDPS.
 DR Pfam: PF00701; DHDPS_1.
 DR PRINTS: PR00146; DHDPSINTHASE.
 DR PRODOM: PD001859; DHDPS_1.
 DR TIGRfam: TIGR00674; dapa; 1.
 DR PROSITE: PS00665; DHDPS_1; 1.
 DR PROSITE: PS00666; DHDPS_2; 1.
 KW Lyase; Diaminopimelate biosynthesis; lysine biosynthesis;
 KW Allosteric enzyme; Chloroplast; Transit peptide; Multigene family.
 FT TRANSIT 1 38 CHLOROPLAST (POTENTIAL)
 FT CHAIN 39 364 DIHYDRODIPICOLINATE SYNTHASE 1.
 FT ACT_SITE 221 221 BY SIMILARITY.
 FT CONFLICT 285 285 A -> E (IN REF. 1).
 SQ SEQUENCE 364 AA; 40464 MW; 73895B17AA79E6EE CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 364;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPN 5
 Db 39 AVPN 43
 RESULT 7
 DAP2_ARATH STANDARD; PRT; 365 AA.
 ID DAP2_ARATH
 AC 09FVC8; 022129;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dihydrodipicolinate synthase 2, chloroplast precursor (EC 4.2.1.52)
 DE (DHDPS 2).
 GN DHDPS2 OR AT2G45440 OR F4L23_5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Wassilewskija;
 RA Sarrobert C., Thibaud M.-C., Contard-David P., Glineste S.,
 RA Bechtold N., Robaglia C., Nussbaum L.,
 RT "Accumulation of threonine resulting from mutation in a new
 RT dihydrodipicolinate synthase gene altered Arabidopsis growth.";
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman J.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.,
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 CC -1- CATALYTIC ACTIVITY: L-aspartate 4-semialdehyde + pyruvate =
 CC dihydrodipicolinate + 2 H(2)O.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 CC semialdehyde; first step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the DHDPS family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

ID	CDCK2_CHERU	STANDARD:	PRT:	294 AA.
AC	PS3101;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DE	28-FEB-2003 (Rel. 41, Last annotation update)			
OS	Cell division control protein 2 homolog (EC 2.7.1.1-) (p34cdc2).			
OC	CDCK2 OR CDK34.			
OC	Chenopodium rubrum (Red goosefoot) (Pigweed).			
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:			
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:			
OC	Caryophyllidae: Caryophyllales: Chenopodiaceae: Chenopodium.			
OX	NCBI_TaxID=3560;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RA	Renz A., Schmeizl B., Beck E.;			
RT	"Nucleotide sequence of a cDNA encoding a CDK34-protein kinase from a			
RL	photoautotrophic cell suspension culture of Chenopodium rubrum L.";			
RL	(in) Plant Gene Register PGR97-045.			
CC	-1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC			
CC	CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES			
CC	THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.			
CC	-1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES			
CC	THE ENZYME. WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY			
CC	SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	CDCK2/CDKX SUPERFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: Y10160; CAA71242.1; -.			
DR	HSSP: P24941; IHCL.			
DR	InterPro: IPR000719; Prot_Kinase.			
DR	InterPro: IPR002290; Ser_thr_pkinase.			
DR	Pfam: PF00069; pkinase; 1.			
DR	ProDom: PD000001; Prot_Kinase; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.			
KW	Transferase; Serine/threonine protein kinase; ATP-binding;			
KW	Cell cycle; Cell division; Mitosis; Phosphorylation.			
FT	DOMAIN 4 287			
FT	NP_BIND 10 18			
FT	BINDING 33 33			
FT	ACT_SITE 127 127			
FT	MOD_RES 14 14			
FT	MOD_RES 15 15			
FT	MOD_RES 161 161			
FT	SEQUENCE 294 AA; 33632 MW; B82748F84B600B6 CRC64;			
QY	Query Match	100.0%;	Score 25;	DB 1; Length 294;
QY	Best Local Similarity	100.0%;	Pred. NO. 67;	
QY	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 AYPVN 5			
QY				
QY	251 AYPVN 255			

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 [1]
 RP SEQUENCE FROM N.A.; SEQUENCE OF 22-41, AND CHARACTERIZATION.
 RC STRAIN=Japanese white; TISSUE=Pancreas, and Utricle;
 RX MEDLINE=97361938; PubMed=9230129;
 RA Yasuda T., Takeshita H., Nakajima T., Hosomi O., Nakashima Y.,
 Kishi K.;
 RT "Rabbit Dnaase I: purification from urine, immunological and
 RT proteochemical characterization, nucleotide sequence, expression in
 RT tissues, relationships with other mammalian Dnaases I and phylogenetic
 RT analysis.";
 RL Biochem. J. 325:465-473(1997).
 CC -1- FUNCTION: AMONG OTHER FUNCTIONS, SEEMS TO BE INVOLVED IN CELL
 CC DEATH BY APOPTOSIS. BINDS SPECIFICALLY TO G-ACTIN AND BLOCKS ACTIN
 CC POLYMERIZATION. PREFERENTIALLY ATTACKS DOUBLE-STRANDED DNA AND
 CC PRODUCES OLIGONUCLEOTIDES WITH 5'-PHOSHO AND 3'-HYDROXY TERMINI.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphodinucleotide and 5'-phosphooligonucleotide end-products.
 CC -1- COFACTOR: DIVALENT CATIONS, PARTICULARLY CALCIUM AND MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: SECRETORY PROTEIN, STORED IN ZYMOGEN
 CC GRANULES AND FOUND IN THE NUCLEAR ENVELOPE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EQUIVALENT LEVELS IN PANCREAS AND PAROTID
 CC GLAND. LOW AMOUNTS IN KIDNEY, LIVER, SMALL INTESTINE, STOMACH AND
 CC THYMUS.
 CC -1- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D82875; BAA21724.1; -;
 DR HSSP; P00639; 2DNT.
 DR InterPro: IPR001582; Dnaase.I.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR PRINTS; PR00130; DNASE1.
 DR PRODOM; PD005408; Dnaase.I.N; 1.
 DR SMART; SM00476; Dnaase1c; 1.
 DR PROSITE; PS00919; DNASE1_1; 1.
 DR PROSITE; PS00918; DNASE1_2; FALSE NEG.
 KW Hydrolyase; Endonuclease; Nuclease; Glycoprotein; Calcium; Signal;
 KW Apoptosis; Actin-binding.
 FT SIGNAL 1 21
 FT CHAIN 1 21
 FT DISULFID 122 125
 FT ACT_SITE 99 99 BY SIMILARITY.
 FT ACT_SITE 155 155 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC...).
 SQ SEQUENCE 281 AA; 31345 MW; 6B6535FC1FE29E8 CMC64;
 Query Match 100.0%; Score 25; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPN 5
 Db 245 AVPN 249
 RESULT 3
 ID DRNL_MOUSE STANDARD; PRT; 284 AA.
 AC P49183; 070532;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Deoxyribonuclease I precursor (EC 3.1.21.1) (Dnaase I);

GN DNASE1 OR DNL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Heart;
 RX MEDLINE=95160718; PubMed=7857306;
 RA Beltsch M.C., Imbler M., French L.E., Tschoop J.;
 RT "Genomic organisation and expression of mouse deoxyribonuclease I.";
 RT Biochem. Biophys. Res. Commun. 207:62-68(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Kidney, and Parotid gland;
 RX MEDLINE=97335420; PubMed=9192086;
 RA Takeshita H., Yasuda T., Nakajima T., Hosomi O., Nakashima Y.,
 Kishi K.;
 RT "Mouse deoxyribonuclease I (Dnaase I): biochemical and immunological
 RT characterization, cDNA structure and tissue distribution.";
 RL Biochem. Mol. Biol. Int. 42:65-75(1997).
 CC -1- FUNCTION: AMONG OTHER FUNCTIONS, SEEMS TO BE INVOLVED IN CELL
 CC DEATH BY APOPTOSIS. BINDS SPECIFICALLY TO G-ACTIN AND BLOCKS
 CC ACTIN POLYMERIZATION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphodinucleotide and 5'-phosphooligonucleotide end-products.
 CC -1- COFACTOR: DIVALENT CATIONS, PARTICULARLY CALCIUM AND MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: SECRETORY PROTEIN, STORED IN ZYMOGEN
 CC GRANULES AND FOUND IN THE NUCLEAR ENVELOPE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN STRIATED MUSCLE, KIDNEY, INTESTINE, LIVER,
 CC LYMPHNODES, BUT NOT IN THE HEART, SPLEEN OR PANCREAS.
 CC -1- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U00478; AAA03710.1; -;
 DR EMBL; D83038; BAA28622.1; -;
 DR PIR; JC2526; JC2526.
 DR HSSP; P00639; 3DNI.
 DR MGD; MGI:103157; Dnaase1.
 DR InterPro: IPR001582; Dnaase.I.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR PRINTS; PR00130; DNASE1.
 DR PRODOM; PD005408; Dnaase.I.N; 1.
 DR SMART; SM00476; Dnaase1c; 1.
 DR PROSITE; PS00918; DNASE1_1; 1.
 DR PROSITE; PS00919; DNASE1_2; 1.
 KW Hydrolyase; Endonuclease; Nuclease; Glycoprotein; Calcium; Signal;
 KW Apoptosis; Actin-binding.
 FT SIGNAL 1 22
 FT CHAIN 1 22
 FT DISULFID 123 126
 FT ACT_SITE 100 100 BY SIMILARITY.
 FT ACT_SITE 156 156 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC...).
 FT CONFLICT 239 240 AG -> VR (IN REF. 1).
 SQ SEQUENCE 284 AA; 32027 MW; 8682E20515EB510 CMC64;
 Query Match 100.0%; Score 25; DB 1; Length 284;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPN 5

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 ; Search time 4.33824 Seconds

(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893d-56

Perfect score: 25

Sequence: 1 AVVPN 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	265	1	VNST_RVFPVZ
2	25	100.0	281	1	DRN1_RABIT
3	25	100.0	284	1	DRN1_MOUSE
4	25	100.0	294	1	CC2A_ANTMA
5	25	100.0	294	1	CCD2_CHERU
6	25	100.0	364	1	DAP1_ARATH
7	25	100.0	365	1	DAP2_ARATH
8	25	100.0	387	1	THIT_MYCPN
9	25	100.0	418	1	CNRC_ALCEU
10	25	100.0	466	1	SYN_VIBPA
11	25	100.0	466	1	SYN_VIBPA
12	25	100.0	466	1	SYN_VIBPA
13	25	100.0	624	1	NIFA_AZOLI
14	25	100.0	625	1	NIFA_AZOB
15	25	100.0	954	1	FLEY_CAVCR
16	25	100.0	964	1	YOKA_SCHPO
17	25	100.0	1360	1	GLI1_XENLA
18	25	100.0	152	1	MLI2_BRARE
19	25	100.0	153	1	MLI1_BRARE
20	25	100.0	154	1	MLIA_PIG
21	25	100.0	169	1	V55_BPT7
22	25	100.0	232	1	6PGL_HAEIN
23	25	100.0	242	1	PEPE_ANASP
24	25	100.0	249	1	Y084_METJA
25	25	100.0	257	1	MLIA_BOVIN
26	25	100.0	283	1	YVAA_BACSU
27	25	100.0	318	1	KPR2_YEAST
28	25	100.0	319	1	KPRS_CIOAB
29	25	100.0	337	1	G3P_METHA
30	25	100.0	338	1	G3P_METHB
31	25	100.0	346	1	G3P_METHO
32	25	100.0	346	1	MLIC_CHICK
33	25	100.0	353	1	MLIA_CHICK

34	24	96.0	366	1	MLIA_SHEEP	P48040	ovis aries
35	24	96.0	379	1	Y528_SYNV3	O5518	synechocyst
36	24	96.0	380	1	GLGC_BACSU	P39122	bacillus su
37	24	96.0	420	1	MLIC_XENLA	P49219	xenopus lae
38	24	96.0	420	1	SECT_GUTH	P28527	guillardi
39	24	96.0	429	1	APEB_PSEAE	O9hy3	pseudomonas
40	24	96.0	431	1	SYH_PYRAB	O9uy31	pyrococcus
41	24	96.0	449	1	MURC_HELPJ	O92112	helicobacte
42	24	96.0	449	1	MURC_HELPJ	O25340	helicobacte
43	24	96.0	484	1	PANF_HAEIN	P44663	haemophilus
44	24	96.0	555	1	FPRB_MICLE	O33064	mycobacteri
45	24	96.0	562	1	SIS2_YEAST	P36024	saccharomyc

ALIGNMENTS

RESULT 1	VNST_RVFPVZ	STANDARD;	PRT;	265 AA.
ID	VNST_RVFPVZ			
AC	P21698;			
DT	01-MAY-1991 (Rel. 18, Last sequence update)			
DT	01-MAY-1991 (Rel. 18, Last sequence update)			
DT	01-AUG-1992 (Rel. 23, Last annotation update)			
DE	Nonstructural protein NS-S.			
OS	Rift valley fever virus (strain ZH-548 M12) (RVFV).			
OC	Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.			
OX	NCBI_TaxID=11589;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=9111992; PubMed=1846496;			
RA	Giorgi C., Accardi L., Nicoletti L., Gro M.C., Takehara K.,			
RA	Hilditch C., Morikawa S., Bishop D.H.L.;			
RT	"Sequences and coding strategies of the S RNAs of Toscana and Rift			
RT	Valley fever viruses compared to those of Punta Toro, Sicilian			
RT	Sandfly fever, and Uukuniemi viruses.";			
RT	Virology 180:738-753(1991).			
RL	-1 SIMILARITY: NS-S FROM PUNTA TORO, RIFT VALLEY FEVER, SANDFLY FEVER			
CC	SICILIAN, TOSCANA, AND UUKUNIEMI VIRUSES ARE EVOLUTIONARY RELATED.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
CC	EMBL: X53771; CAA37788.1; -			
DR	PIR: C38552; MNUVRV.			
DR	PIR: C38552; MNUVRV.			
KW	Nonstructural protein.			
SO	SEQUENCE 265 AA; 29903 MW; EF2608D29CEB6C84 CRC64;			
QY	1 AVVPN 5			
Db	208 AVVPN 212			
Query Match	100.0%; Score 25; DB 1; Length 265;			
Best Local Similarity	100.0%; Pred. No. 60;			
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
RESULT 2	DRN1_RABIT	STANDARD;	PRT;	281 AA.
ID	DRN1_RABIT			
AC	O18998;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Deoxyribonuclease I precursor (EC 3.1.21.1) (DNase I).			
GN	DNASE1 OR DNLI.			
OS	Oryctolagus cuniculus (Rabbit).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

A:Cross-references: GB:AE003924; GB:AE03849; NID:95105750; PIDN:AAF83658.1; GSPDB:GN004
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H
 as-Veto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitaajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savaas
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0848

Query Match 100.0%; Score 25; DB 2; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 159 AVPN 163

RESULT 13

C72590
 Probable hexosyltransferase (EC 2.4.1.-) APE1191 [similarity] - Aeropyrum pernix (strain
 N:Alternate names: Probable capM protein
 C:Species: Aeropyrum pernix
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000
 C:Accession: C72590
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jinno, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: C72590
 A:Molecule type: DNA
 A:Residues: 1-363 <KAM>
 A:Cross-references: DDBJ:AP000061; NID:95104821; PIDN:BA80177.1; PID:95104863
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1191
 C:Superfamily: probable hexosyltransferase YFXN
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 100.0%; Score 25; DB 1; Length 363;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 166 AVPN 170

RESULT 14

S46304
 dihydrodipicolinate synthase (EC 4.2.1.52) precursor [similarity] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S46304
 R:Vauterin, M.; Jacobs, M.
 Plant Mol. Biol. 25, 545-550, 1994
 A:Title: Isolation of a poplar and an Arabidopsis thaliana dihydrodipicolinate synthase
 A:Reference number: S46304; MUID:94325478; PMID:8049377
 A:Accession: S46304
 A:Molecule type: mRNA
 A:Residues: 1-364 <VAU>
 A:Cross-references: EMBL:X72971; NID:9530004; PIDN:CAB45642.1; PID:95139329

C:Genetics:
 A:Genome: nuclear
 A:Function:
 A:Description: catalyzes the formation of dihydrodipicolinate from L-aspartate 4-semi
 A:Pathway: lysine biosynthesis
 A:Note: first step in diaminopimelate and lysine biosynthesis; sensitive to lysine in
 C:Superfamily: dihydrodipicolinate synthase
 C:Keywords: carbon-oxygen lyase; chloroplast; hydro-lyase; lysine biosynthesis
 F:1-38/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:39-364/Product: dihydrodipicolinate synthase #status predicted <MAT>
 F:221/Active site: Lys (covalent pyruvate-binding) #status predicted

Query Match 100.0%; Score 25; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 39 AVPN 43

RESULT 15

T47899
 dihydrodipicolinate synthase precursor - Arabidopsis thaliana
 N:Alternate names: protein T4C21.290
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
 C:Accession: T47899
 R:Cholme, N.; Robert, C.; Brothier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.;
 W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: Z24479
 A:Accession: T47899
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <CHO>
 A:Cross-references: EMBL:AL162295
 A:Experimental source: cultivar Columbia; BAC clone T4C21
 C:Genetics:
 A:Map position: 3
 A:Introns: 29/2; 56/2
 A:Note: T4C21.290
 C:Superfamily: dihydrodipicolinate synthase

Query Match 100.0%; Score 25; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 39 AVPN 43

Search completed: September 4, 2003, 21:10:47
 Job time : 9.79412 secs

A:Gene: lin2335

Query Match

100.0%; Score 25; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5
 |||||
 Db 239 AAVPN 243

RESULT 8

transcription regulators (lysr family) homolog lmo2233 [imported] - *Listeria monocytogenes*
 C:Species: *Listeria monocytogenes*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: A11353

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bleecker, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurupkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A.; Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11353
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <GUA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00311.1; PID:g16411703; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo2233

Query Match

100.0%; Score 25; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5
 |||||
 Db 239 AAVPN 243

RESULT 9

T17115
 protein kinase cdc2a (EC 2.7.1.-), cyclin-dependent - garden snaddragon
 C:Species: *Antirrhinum majus* (garden snaddragon)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T17115

R:Robert, P.R.; Gaudin, V.; Lunness, P.; Coen, E.S.; Doonan, J.H.
 Plant Cell 8, 1465-1476, 1996
 A:Title: Distinct classes of cdc2-related genes are differentially expressed during the
 A:Reference number: Z18695; MUID:9643455; PMID:8837502
 A:Accession: T17115
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-302 <FOR>
 A:Cross-references: EMBL:X97637; NID:g1321671; PIDN:CAA66233.1; PID:g1321672
 C:Genetics:
 A:Note: cdc2a
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: phosphotransferase

Query Match

100.0%; Score 25; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5
 |||||
 Db 259 AAVPN 263

RESULT 10

AG2172

hypothetical protein all2934 [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp. PCC 7120
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AG2172

R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, F.; Sasamoto, S.; Watanabe, A.; Irigu, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG2172
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA074633.1; PID:g17132028; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all2934
 C:Superfamily: *Escherichia coli* glutaminase homolog yneH

Query Match

100.0%; Score 25; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5
 |||||
 Db 277 AAVPN 281

RESULT 11

E86794

hypothetical protein ynfJ [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL140
 C:Species: *Lactococcus lactis* subsp. *lactis*
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: E86794

R:Boletini, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Eh
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
 A:Reference number: AB6625; MUID:21235186; PMID:11337471
 A:Accession: E86794
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <STO>
 A:Cross-references: GB:AE005176; PID:g12724340; PIDN:AK05455.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ynfJ

Query Match

100.0%; Score 25; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVPN 5
 |||||
 Db 249 AAVPN 253

RESULT 12

C82735

conserved hypothetical protein XF0848 [imported] - *Xylella fastidiosa* (strain 9a5c)

C:Species: *Xylella fastidiosa*
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: C82735
 R:Anonymous. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A:Reference number: AB2515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: C82735
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-351 <SIM>

A:Contents: annotation
C:Genetics:
A:Gene: SMD20131
A:Genome: plasmid
C:Superfamily: isoguinoline-1-oxidoreductase alpha chain; ferredoxin [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein
F:41.46.49.61/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 100.0%; Score 25; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
|||||
DB 17 AVPN 21

RESULT 3
F75078
hypothetical protein PAB1691 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000
C:Accession: F75078
R:Anonymous; Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75078
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <KAM>
A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB49931.1; PID:g545844
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1691
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1869

Query Match 100.0%; Score 25; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
|||||
DB 31 AVPN 35

RESULT 4
H59433
D1100H3.4 protein [imported] - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 23-Sep-2002
C:Accession: H59433
R:Smith, M.
submitted to Genbank, March 2001
A:Reference number: H59433
A:Accession: H59433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <SMI>
A:Cross-references: GB:CA017688; PID:g13397832; PIDN:CA017688.2

Query Match 100.0%; Score 25; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
|||||
DB 142 AVPN 146

RESULT 5
MN009
nonstructural protein NS - Rift Valley fever virus

C:Species: Rift Valley fever virus
A:Note: host (mosquito); Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C:Accession: C38552; S11743
R:Giorgio, C.; Accardi, L.; Nicoletti, L.; Gro, M.C.; Takehara, K.; Hilditch, C.; Mori
Virology 180, 738-753, 1991
A:Title: Sequences and coding strategies of the 5 RNAs of Toscana and Rift Valley fever
A:Reference number: A38552; MUID:91111992; PMID:1846496
A:Accession: C38552
A:Molecule type: genomic RNA
A:Residues: 1-265 <GID>
A:Cross-references: GB:X53771; NID:g61928; PIDN:CA37788.1; PID:g61929
C:Genetics:
A:Map position: segment S
C:Superfamily: Punta Toro virus nonstructural protein NS
C:Keywords: nonstructural protein

Query Match 100.0%; Score 25; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
|||||
DB 208 AVPN 212

RESULT 6
JC2526
deoxyribonuclease I (EC 3.1.21.1) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JC2526
R:Peitsch, M.C.; Immler, M.; French, L.E.; Tschoopp, J.
Biochem. Biophys. Res. Commun. 207, 62-68, 1995
A:Title: Genomic organisation and expression of mouse deoxyribonuclease I.
A:Reference number: JC2526; MUID:95160718; PMID:7857306
A:Accession: JC2526
A:Molecule type: mRNA
A:Residues: 1-284 <PEI>
A:Cross-references: EMBL:U00478; NID:g437052; PIDN:AA03710.1; PID:g437053
C:Superfamily: deoxyribonuclease I
C:Keywords: actin binding; calcium binding; hydrolase

Query Match 100.0%; Score 25; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
|||||
DB 246 AVPN 250

RESULT 7
AB1724
transcription regulators (LysR family) homolog lin235 [imported] - Listeria innocua
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1724
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kunz, M.; Kunst, F.; Kurupkat, G.; Madeno, E.; Maltournam, A.;
Ok, C.; Schlueter, T.; Simoes, N.; Tjierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1724
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <GLA>
A:Cross-references: GB:AL592022; PIDN:CA037562.1; PID:g16414858; GSPDB:GN00178
A:Experimental source: strain C1p11262
C:Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 7.79412 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-56

Perfect score: 25

Sequence: 1 AVVPN 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR-76:*
2: pir1:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	149	2	regulatory protein
2	25	100.0	161	2	probable oxidoredu
3	25	100.0	229	2	hypothetical prote
4	25	100.0	242	2	DJ1100H13.4 protei
5	25	100.0	265	1	nonstructural prot
6	25	100.0	284	1	deoxyribonuclease
7	25	100.0	292	2	transcription regu
8	25	100.0	292	2	transcription regu
9	25	100.0	302	2	protein kinase cdc
10	25	100.0	334	2	hypothetical prote
11	25	100.0	334	2	hypothetical prote
12	25	100.0	351	2	conserved hypotet
13	25	100.0	363	1	probable hexosyltr
14	25	100.0	364	2	dihydrodipicolinat
15	25	100.0	364	2	dihydrodipicolinat
16	25	100.0	365	2	probable hexosyltr
17	25	100.0	387	1	probable dihydrodi
18	25	100.0	408	2	M6372 homolog G12
19	25	100.0	418	2	hypothetical prote
20	25	100.0	442	2	nickel-cobalt resi
21	25	100.0	442	2	YOPN [imported] -
22	25	100.0	467	2	type III secretion
23	25	100.0	467	2	hypothetical prote
24	25	100.0	468	2	hypothetical prote
25	25	100.0	481	2	asparaginyl-tRNA s
26	25	100.0	503	2	probable membrane
27	25	100.0	513	2	hypothetical prote
28	25	100.0	589	2	hypothetical prote
29	25	100.0	604	2	probable splicing
			614	2	phaseolin G-box bi
			614	2	phaseolin G-box bi

30	25	100.0	620	2	yeast nrd1-like pr
31	25	100.0	623	2	MYC-related DNA bi
32	25	100.0	624	2	regulatory protein
33	25	100.0	625	2	regulatory protein
34	25	100.0	642	2	phaseolin G-box bi
35	25	100.0	663	2	cytochrome O ubiq
36	25	100.0	693	2	4-alpha-glucanotri
37	25	100.0	735	2	cation transport A
38	25	100.0	850	2	hypothetical prote
39	25	100.0	954	2	regulatory protein
40	25	100.0	964	2	hypothetical prote
41	25	100.0	1231	2	probable nitrate r
42	25	100.0	1360	2	DNA binding protei
43	24	96.0	65	2	hypothetical prote
44	24	96.0	68	2	hypothetical prote
45	24	96.0	72	2	potassium channel

ALIGNMENTS

RESULT 1

S02164

regulatory protein flae - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 04-Mar-2000

C:Accession: S02164

R:Kaplan, J.B.; Dingwall, A.; Bryan, R.; Champier, R.; Shapiro, L.

J. Mol. Biol. 205, 71-83, 1989

A:Title: Temporal regulation and overlap organization of two Caulobacter flagellar ge

A:Reference number: S02164; MUID:89178645; PMID:2648000

A:Accession: S02164

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-149 <KAP>

A:Genetics:

A:Gene: flae

C:Superfamily: Caulobacter crescentus regulatory protein flae

C:Keywords: transcription regulation

Query Match	Score	DB 2:	Length
Best Local Similarity	100.0%	Pred. No. 57;	
Matches	5;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1	AVVPN 5	
DB	47	AVVPN 51	

RESULT 2

C95858

probable oxidoreductase protein SMD20131 [imported] - Sinorhizobium meliloti (strain

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 22-Oct-2001

C:Accession: C95858

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing e

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: C95858

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <KUR>

A:Cross-references: GB:A1591985; PIDN:CAC48531.1; PID:915140003; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pel, A.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

l.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

PN W0200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
DR
XX WPI: 2001-488900/53.
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 30716; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 27 AA:

Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 |||||
DB 18 AVPN 22

RESULT 15
AAM18247
ID AAM18247 standard; Protein: 27 AA.
XX
AC AAM18247;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #4681 encoded by probe for measuring cervical gene expression.
XX
KW Probe: human: microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
OS
PN W0200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488901/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 23073; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SNP: see A110068-A128459). The present sequence is a peptide encoded
CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 27 AA:

Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 |||||
DB 18 AVPN 22

Search completed: September 4, 2003, 21:00:29
Job time : 26.7353 secs

OY 1 AVVPN 5
 DB 18 AVVPN 22

RESULT 12

ABR22567 standard; Protein: 27 AA.

AC ABR22567;

DT 23-JAN-2002 (first entry)

DE Protein #4566 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

PS Claim 15; SEQ ID NO 24337; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG4748-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAVPN 5
| | | | |
Db 18 AAVPN 22
RESULT 10
ABB32028
ID ABB32028 standard; Peptide: 27 AA.
XX
AC ABB32028;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #4679 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes
XX
PS Claim 27; SEQ ID NO 14996; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias

CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAVPN 5
| | | | |
Db 18 AAVPN 22
RESULT 11
ABB37278
ID ABB37278 standard; Peptide: 27 AA.
XX
AC ABB37278;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #4784 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 29913; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HPP), VVPN derived peptide.

XX Sequence 8 AA:

Query Match 100.0%; Score 25; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5

Db 3 AVVPN 7

RESULT 8

AAE16730

ID AAE16730 standard; peptide; 8 AA.

XX AAE16730;

DT 09-APR-2002 (first entry)

DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.

XX Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KM vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KM hypotensive; chronic obstructive pulmonary disease; arteriosclerosis;
 KM restenosis post-angioplasty; antidiabetic; anticoagulant; thrombolytic;
 KM cyclic.

XX Ligamentum nuchae.
 OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note="This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"

FT Modified-site 8 /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"

XX WO200191700-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17384.

XX 30-MAY-2000; 2000US-0580110.

XX 30-MAY-2000; 2000US-0580156.

XX 30-MAY-2000; 2000US-0580893.

XX 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING L.L.C.

XX Miltis TF, Sandberg LB, Jimenez F;

XX WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide
 XX substances which mimic the sequences found in elastin
 XX Claim 24; Page 21; 53pp; English.

CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HPP), VVPN derived cyclic peptide.

XX Sequence 8 AA:

Query Match 100.0%; Score 25; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5

Db 3 AVVPN 7

RESULT 9

ABG52100

ID ABG52100 standard; Peptide; 27 AA.

XX ABG52100;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 30748.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KM hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 30748; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/ fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult
 XX liver. (I) may be used for predicting, measuring and displaying gene

CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 7 AA:
Query Match 100.0%; Score 25; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVPN 5
| | | | |
DB 2 AVPN 6
RESULT 6
ID AAE16729 standard; peptide: 7 AA.
XX AAE16729;
XX 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #9.
XX
XX Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
KW cyclic.
XX
XX Ligamentum nuchae.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note="This residue forms a cyclic structure with Cys
FT at position 7 with copper as a chelating agent"
FT Modified-site 7
FT /note="This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent"
XX
XX WO200191700-A2.
XX
XX 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17384.
XX
XX 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
XX (CONN-) CONNECTIVE TISSUE IMAGING ENGINEERING LLC.
XX
XX Mitsu TF, Sandberg LB, Jimenez F;
XX
XX WPI: 2002-106259/14.
XX
XX Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.

XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived cyclic peptide.
XX
SQ Sequence 7 AA:
Query Match 100.0%; Score 25; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVPN 5
| | | | |
DB 2 AVPN 6
RESULT 7
ID AAE16727 standard; peptide: 8 AA.
XX AAE16727;
XX 09-APR-2002 (first entry)
XX
XX Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
XX
XX Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
KW
XX
XX Ligamentum nuchae.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note="This residue forms a cyclic structure with Cys
FT at position 7 with copper as a chelating agent"
FT Modified-site 7
FT /note="This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent"
XX
XX WO200191700-A2.
XX
XX 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17384.
XX
XX 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
XX (CONN-) CONNECTIVE TISSUE IMAGING ENGINEERING LLC.
XX
XX Mitsu TF, Sandberg LB, Jimenez F;
XX
XX WPI: 2002-106259/14.
XX
XX Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.
XX
XX The present invention relates to a composition for the treatment of
XX mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
XX their biological equivalents. The composition is useful not only for
XX improving the cosmetic appearance of skin, but for the treatment of

XX Claim 24; Page 21; 53pp; English.

PS The present invention relates to a composition for the treatment of
XX mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

XX Sequence 6 AA;

Query Match 100.0%; Score 25; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
|||||
Db 2 AVVPN 6

RESULT 4
AAE16724 standard; peptide; 6 AA.

XX AAE16724;

XX 09-APR-2002 (first entry)

XX Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #4.

XX Hydrolysed elastin peptide; HEP, therapy; cosmetic appearance; cardiant;
XX vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
XX coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
XX hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
XX restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.

XX Ligamentum nuchae.
XX Synthetic.

XX Key Location/Qualifiers
XX Modified-site 6 /note="C-terminal amide"

XX WO200191700-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17384.

XX 30-MAY-2000; 2000US-0580110.

XX 30-MAY-2000; 2000US-0580156.

XX 30-MAY-2000; 2000US-0580893.

XX 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.

XX Mlts TF, Sandberg LB, Jimenez F;

XX WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin

PS Claim 24; Page 21; 53pp; English.

XX The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

XX Sequence 6 AA;

Query Match 100.0%; Score 25; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
|||||
Db 2 AVVPN 6

RESULT 5
AAE16726 standard; peptide; 7 AA.

XX AAE16726;

XX 09-APR-2002 (first entry)

XX Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #6.

XX Hydrolysed elastin peptide; HEP, therapy; cosmetic appearance; cardiant;
XX vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
XX coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
XX hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
XX restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.

XX Ligamentum nuchae.
XX Synthetic.

XX Key Location/Qualifiers
XX Modified-site 6 /note="C-terminal amide"

XX WO200191700-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17384.

XX 30-MAY-2000; 2000US-0580110.

XX 30-MAY-2000; 2000US-0580156.

XX 30-MAY-2000; 2000US-0580893.

XX 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.

XX Mlts TF, Sandberg LB, Jimenez F;

XX WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin

PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 5 AA:
Query Match 100.0%; Score 25; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
DB 1 AVVPN 5
RESULT 2
ID AAE16723 standard; peptide: 5 AA.
XX
AC AAE16723;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #3.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 5 /note="C-terminal amide"
FT
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;

XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 5 AA:
Query Match 100.0%; Score 25; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
DB 1 AVVPN 5
RESULT 3
ID AAE16722 standard; peptide: 6 AA.
XX
AC AAE16722;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #2.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin -

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 : Search time 25.7353 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893D-56

Perfect score: 25

Sequence: 1 AAVPN 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: A.Geneseq-19jun03:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

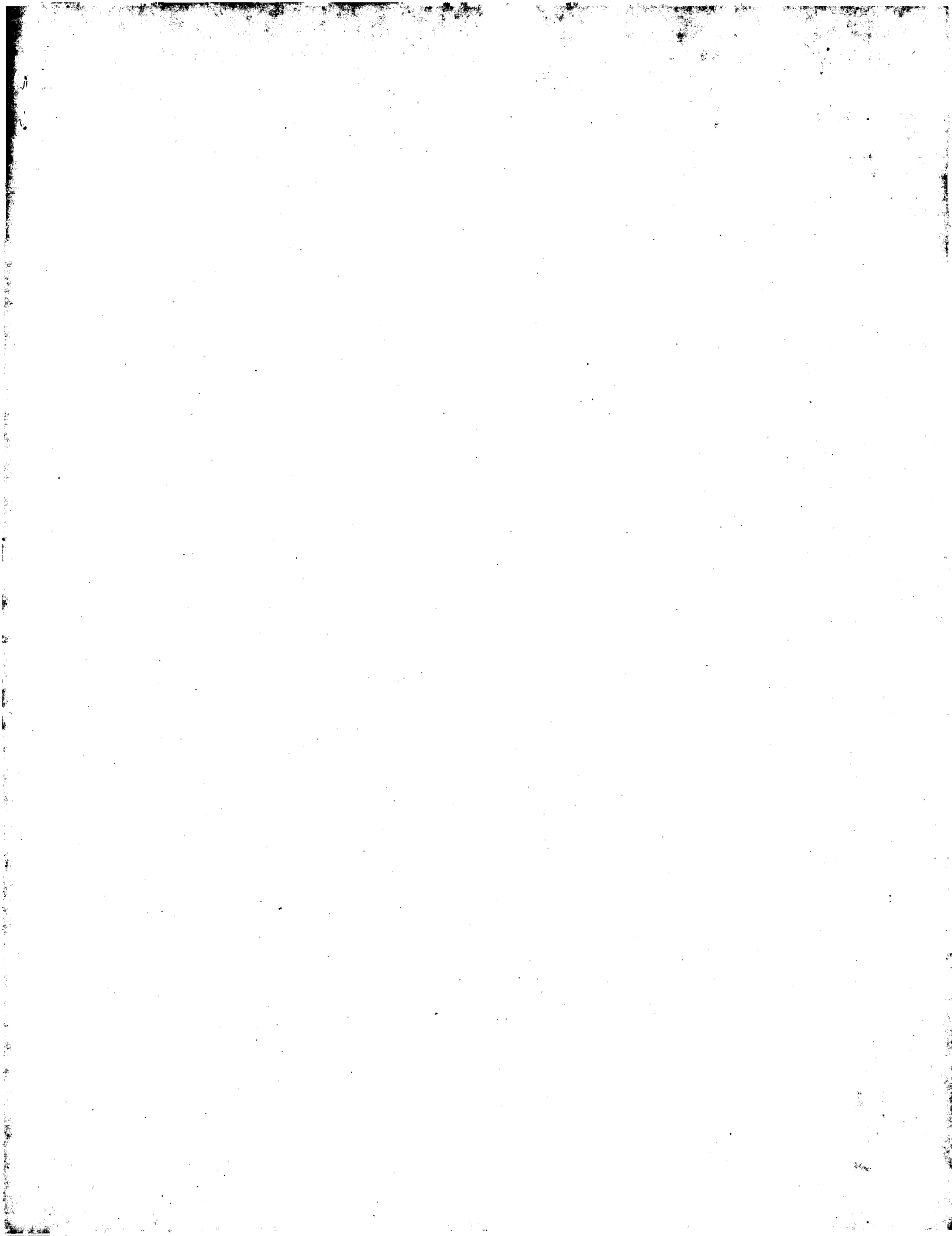
Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	5	23	AAE16721
2	25	100.0	5	23	AAE16723
3	25	100.0	6	23	AAE16722
4	25	100.0	6	23	AAE16724
5	25	100.0	7	23	AAE16726
6	25	100.0	7	23	AAE16729
7	25	100.0	8	23	AAE16727
8	25	100.0	8	23	AAE16730
9	25	100.0	27	22	ABG52100

10	25	100.0	27	22	ABB32028
11	25	100.0	27	22	ABB37278
12	25	100.0	27	22	ABB25678
13	25	100.0	27	22	AAW57980
14	25	100.0	27	22	AAW70410
15	25	100.0	27	22	AAW18247
16	25	100.0	27	22	AAW30740
17	25	100.0	27	22	AAW05859
18	25	100.0	27	22	ABG40049
19	25	100.0	57	22	AAU66327
20	25	100.0	80	22	AAU21097
21	25	100.0	91	22	AAU67514
22	25	100.0	107	22	AAU44534
23	25	100.0	110	23	ABG30982
24	25	100.0	112	21	AAW33017
25	25	100.0	138	24	ABP79356
26	25	100.0	188	21	AAW32721
27	25	100.0	193	21	AAW33083
28	25	100.0	207	21	AAW32670
29	25	100.0	217	20	AAW32095
30	25	100.0	235	21	AAW32764
31	25	100.0	259	22	AAU68540
32	25	100.0	292	23	ABW48286
33	25	100.0	334	23	ABW54696
34	25	100.0	354	21	AAW40467
35	25	100.0	364	21	AAW24692
36	25	100.0	364	23	ABW92747
37	25	100.0	365	21	AAW40466
38	25	100.0	365	20	ABW92129
39	25	100.0	442	20	AAW35333
40	25	100.0	442	23	ABW56014
41	25	100.0	442	23	ABW98223
42	25	100.0	442	24	ABW66279
43	25	100.0	455	22	AAW92773
44	25	100.0	467	24	AAW30065
45	25	100.0	485	20	AAW45267

ALIGNMENTS

RESULT 1	
AAE16721	AAE16721 standard; peptide: 5 AA.
ID	AAE16721 standard; peptide: 5 AA.
AC	AAE16721:
XX	
XX	
XX	09-APR-2002 (first entry)
XX	
XX	Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #1.
XX	
XX	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
XX	KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
XX	KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
XX	KW hypotensive; chronic obstructive pulmonary disease; antidiabetic; antidiabetic;
XX	restenosis post-angioplasty; antidiabetic; antidiabetic; thrombolytic.
XX	
XX	Ligamentum nuchae.
XX	
XX	Synthetic.
XX	
XX	WO200191700-A2.
XX	
XX	06-DEC-2001.
XX	
XX	30-MAY-2001; 2001WO-US17384.
XX	
XX	30-MAY-2000; 2000US-0580110.
XX	PR 30-MAY-2000; 2000US-0580156.
XX	PR 30-MAY-2000; 2000US-0580893.
XX	PR 30-MAY-2000; 2000US-0584001.
XX	
XX	(CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX	

Peptide #4679 enco
Peptide #4784 enco
Protein #4566 enco
Human brain expres
Human bone marrow
Peptide #4681 enco
Peptide #4777 enco
Peptide #4541 enco
Human peptide enco
Protonibacterium
Human novel foetal
Protonibacterium
Protonibacterium
Human macroprotein
Pinus radiata tran
N. gonorrhoeae ami
Eucahyplus grandis
Pinus radiata tran
Eucahyplus grandis
Aspen homeobox pro
Eucahyplus grandis
Human novel cytoxi
Listeria monocytog
Lactococcus lactis
Arabidopsis thalia
Arabidopsis thalia
Herbicidally activ
Arabidopsis thalia
Herbicidally activ
Chlamydia pneumoni
Chlamydia polypept
C. psittaci protei
C. glutamicum prote
Arabidopsis thalia
Human secreted pro



```

RESULT 14
RT 08V5M7
ID 08V5M7 PRELIMINARY; PRT; 87 AA.
AC 08V5M7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 08V20.
OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
ON NCBI_TaxID=10468;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.F.,
RA Presnail J.K., Herrmann R., Dolan M., Tingey S., Hu Z.-H., Vlak J.M.;
RT "Genome sequence analysis of Helicoverpa zea single nucleocapsid
RT nucleopolyhedrovirus.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF334030; AAL56165.1;
SQ SEQUENCE 87 AA; 9353 MW; 30A98A48B22E9FEA CRC64;

Query Match 100.0%; Score 21; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPM 4
   ||||
Db 75 VVPM 78

RESULT 15
ID 099H26 PRELIMINARY; PRT; 87 AA.
AC 099H26;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE p10.
OS Helicoverpa armigera nucleopolyhedrovirus G4.
OS Helicoverpa armigera nuclear polyhedrosis virus, and
OS Helicoverpa armigera single nucleocapsid polyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
ON NCBI_TaxID=148363; 51313, 160266;
RN [1]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;
RT "Sequence analysis of the gp37 gene of Heliothis armigera single-
RT nucleocapsid nucleopolyhedrovirus.";
RL Zhongguo Bingduxue 15:35-42(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Wang H., Hu Z., Sun X., Vlak J.M., Chen X.;
RT "Sequence analysis of the iap3 gene of Heliothis armigera single-
RT nucleocapsid nucleopolyhedrovirus.";
RL Zhongguo Bingduxue 15:43-49(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RX MEDLINE=21078302; PubMed=11210934;
RA Wang H., Chen X., Wang H., Arif B.M., Vlak J.M., Hu Z.;
RT "Nucleotide sequence and transcriptional analysis of a putative basic
RT DNA-binding protein of Helicoverpa armigera polyhedrovirus.";
RL Virus Genes 22:113-120(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RX MEDLINE=21064569; PubMed=11125177;
RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;

```

```

RT "The sequence of the Helicoverpa armigera single-nucleocapsid
RT nucleopolyhedrovirus genome.";
RL J. Gen. Virol. 82:241-257(2001).
RN [5]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=CI;
RX PubMed=12050807;
RA Zhang C.X., Wu J.C.;
RT "Genome structure and the p10 gene of the Helicoverpa armigera
RT nucleopolyhedrovirus.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).
RN [7]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=CI;
RA Zhang C.X., Jin W.R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA SPECIES=Helicoverpa armigera single nucleocapsid polyhedrovirus;
RA Wang H., Chen X., Vlak J.M., Hu Z.;
RT "Sequence and transcriptional analysis of the p10 gene of Helicoverpa
RT armigera single nucleocapsid polyhedrovirus.";
RL Ping Tu Hsueh Pao 17:81-86(2001).
DR EMBL: AF271059; AAG53764.1;
DR EMBL: AF303045; AAK96274.1;
DR EMBL: AF265354; AAK57874.1;
SQ SEQUENCE 87 AA; 9331 MW; 30A98644B22293FA CRC64;

QY 1 VVPM 4
   ||||
Db 75 VVPM 78

```

Search completed: September 4, 2003, 21:06:35
 Job time : 20.5027 secs

DE Conserved hypothetical protein.
 GN SAG0550.
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547;
 RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
 RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegli F., Malone D.,
 RA Rinaudo D., Rappunli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL: AE014217; AAM99451.1; -
 DR TIGR: SAG0550; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 74 AA; 7907 MW; C817C7E354630905 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 74;
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVPN 4
 ||||
 Db 7 VVPN 10

RESULT 11
 Q85348 PRELIMINARY; PRT; 80 AA.
 ID Q85348;
 AC Q85348;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Unspecified ORF (Fragment).
 OS Vaccinia virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89094985; PubMed=2783466;
 RA Kotwal G.J., Moss B.;
 RT "Vaccinia virus encodes two proteins that are structurally related to
 RT members of the plasma serine protease inhibitor superfamily.";
 RL J. Virol. 63:600-606(1989).
 DR EMBL: M24217; AAA48342.1; -
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank; 1.
 DR SMART: SM00248; ANK; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 1.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 KW ANK repeat; Repeat.
 FT NON_TER 1
 SQ SEQUENCE 80 AA; 9366 MW; CD9AF4E734EBD132 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 80;
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVPN 4
 ||||
 Db 68 VVPN 71

RESULT 12
 Q93K02 PRELIMINARY; PRT; 81 AA.
 ID Q93K02;
 AC Q93K02;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative replication protein RepC.
 GN REPC.
 OS Yersinia enterocolitica.
 OG Plasmid pYve8081.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8081;
 RX MEDLINE=21295118; PubMed=11402007;
 RA Snellings N.J., Popek M., Lindler L.E.;
 RT "Complete DNA Sequence of Yersinia enterocolitica Serotype O:8 Low-
 RT Calcium-Response Plasmid Reveals a New Virulence Plasmid-Associated
 RT Replicon.";
 RL Infect. Immun. 69:4627-4638(2001).
 DR EMBL: AF336309; AAK69258.1; -
 KW Plasmid.
 SO SEQUENCE 81 AA; 9457 MW; BFE809E656B3FEFF CRC64;

Query Match 100.0%; Score 21; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 7.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVPN 4
 ||||
 Db 40 VVPN 43

RESULT 13
 Q9LA63 PRELIMINARY; PRT; 83 AA.
 ID Q9LA63;
 AC Q9LA63;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Protein P4 (Fragment).
 GN NIKS.
 OS Streptomyces tendae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tue901;
 RX MEDLINE=20177695; PubMed=10712601;
 RA Lauer B., Russwurm R., Bormann C.;
 RT "Molecular characterization of two genes from Streptomyces tendae
 RT Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one
 RT containing nucleoside moiety of the peptidyl nucleoside antibiotic
 RT nikomycin.";
 RL Eur. J. Biochem. 267:1698-1706(2000).
 DR EMBL: AJ250199; CAB75341.1; -
 FT NON_TER 83
 SQ SEQUENCE 83 AA; 9583 MW; 3A441D9A2889EBD2 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVPN 4
 ||||
 Db 59 VVPN 62

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein RC0806.
 GN RC0806.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 DR EMBL; AE008636; AL03344.1; -;
 DR InterPro; IPR002423; AA/rel_permease2.
 DR InterPro; IPR002091; ARA_permease.
 DR PRINTS; PR00166; AROAAPMEASE.
 KW Hypothetical protein: Complete proteome.
 SO SEQUENCE 65 AA; 7177 MW; 8DC3D69FE39D86C CRC64;

Query Match 100.0%; Score 21; DB 16; Length 65;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 IIII
 DB 30 VVPN 33

RESULT 7

O9FTH1 PRELIMINARY; PRT; 66 AA.
 ID O9FTH1
 AC O9FTH1
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE P0410E01.29 protein.
 GN P0410E01.29.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0410E01.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002866; BAB17108.1; -;
 DR HSSP; P01053; IC04.
 DR Gramene; O9FTH1; -;
 DR InterPro; IPR000864; Potato_inhib1.
 DR Pfam; PF00280; potato_inhib1; 1.
 DR PRODOM; PD002604; Potato_inhib1; 1.
 SO SEQUENCE 66 AA; 7204 MW; 165A73BC07B5E0DC CRC64;

Query Match 100.0%; Score 21; DB 10; Length 66;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 IIII
 DB 47 VVPN 50

RESULT 8

O8LRE6 PRELIMINARY; PRT; 69 AA.
 ID O8LRE6
 AC O8LRE6
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE B1015E06.14 protein.
 GN B1015E06.14.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
 RT clone:B1015E06.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003197; BAB92213.1; -;
 DR Gramene; O8LRE6; -;
 SO SEQUENCE 69 AA; 7512 MW; 8D1A1D0373C074A CRC64;

Query Match 100.0%; Score 21; DB 10; Length 69;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 IIII
 DB 66 VVPN 69

RESULT 9

O8H533 PRELIMINARY; PRT; 70 AA.
 ID O8H533
 AC O8H533
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE OJ1753_E03.7 protein.
 GN OJ1753_E03.7.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
 RT clone:OJ1753_E03.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003849; BAC15522.1; -;
 SO SEQUENCE 70 AA; 7779 MW; B575FA41098326C88 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 70;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 IIII
 DB 24 VVPN 27

RESULT 10

O8E115 PRELIMINARY; PRT; 74 AA.
 ID O8E115
 AC O8E115
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)


```
RESULT 2
086596 PRELIMINARY; PRT; 42 AA.
ID 086596
AC 086596;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Polymerase (Fragment).
GN L.
OS Mokoia virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssaviruses.
OX NCBI_TaxID=12538;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93242784; PubMed=8386891;
RA Bourhy H., Kissi B., Tordo N.;
RT "Molecular diversity of the Lyssavirus genus.";
RL Virology 194:70-81(1993).
DR EMBL; S59448; AAB26297.1; -.
FT NON_TER
SQ SEQUENCE 42 AA; 4762 MW; 32E2802AEFDBD878 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVFN 4
Db 26 VVFN 29

RESULT 3
08VJ42 PRELIMINARY; PRT; 44 AA.
ID 08VJ42
AC 08VJ42;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein MT3362.
GN MT3362.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007146; AAN47703.1; -.
DR TIGR; MT3362; -.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 4525 MW; E6429ABC654C276 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 44;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVFN 4
Db 37 VVFN 40

RESULT 4
08F537 PRELIMINARY; PRT; 44 AA.
ID 08F537
```

```
AC 08F537;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN IAI850.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011359; AAN49049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 44 AA; 5447 MW; 0583570EE598A0E8 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 44;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVFN 4
Db 19 VVFN 22

RESULT 5
08EJC4 PRELIMINARY; PRT; 58 AA.
ID 08EJC4
AC 08EJC4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN S00543.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsaplin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vanatrevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nelson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015501; AAN53624.1; -.
DR TIGR; S00543; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6755 MW; 22BCF875342556DA CRC64;

Query Match 100.0%; Score 21; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVFN 4
Db 7 VVFN 10

RESULT 6
092HG5 PRELIMINARY; PRT; 65 AA.
ID 092HG5
AC 092HG5;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 16.4118 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: us-09-580-893d-55

Perfect score: 21

Sequence: 1 VVPN 4

Scoring table:

BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriophage:*
17: sp._archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	31	16	09X142 thermotoga
2	21	100.0	42	12	086596 mokola viru
3	21	100.0	44	16	08YJ42 mycobacteri
4	21	100.0	44	16	08YJ42 mycobacteri
5	21	100.0	58	16	08EJC4 shewanella
6	21	100.0	65	16	09ZHG5 rickettsia
7	21	100.0	66	10	09FTH1 oryza sativ
8	21	100.0	69	10	08IRE6 oryza sativ
9	21	100.0	70	10	08H533 oryza sativ
10	21	100.0	74	16	08E115 streptococc
11	21	100.0	80	12	08E115 streptococc
12	21	100.0	81	2	093KO2 vaccinia vi
13	21	100.0	83	2	09L463 streptomyce
14	21	100.0	87	12	08V5W7 helicoverpa
15	21	100.0	87	12	099H26 helicoverpa
16	21	100.0	87	16	08F2H2 leptospira

17	21	100.0	87	16	08D317 wigglewort
18	21	100.0	91	16	08XPW3 rajstonia s
19	21	100.0	93	16	098RW8 guillardia
20	21	100.0	93	16	098CS4 rhizobium 1
21	21	100.0	101	11	09R136 meriones un
22	21	100.0	102	17	096ZQ7 sulfolobus
23	21	100.0	103	2	053132 rhodobacter
24	21	100.0	105	12	098584 parametium
25	21	100.0	108	15	072902 human immun
26	21	100.0	113	2	056884 yersinia en
27	21	100.0	113	11	09D1Y3 mus musculu
28	21	100.0	116	3	000585 candida par
29	21	100.0	116	5	09NH23 jamides ate
30	21	100.0	116	5	09NH27 spalgis epi
31	21	100.0	116	5	09NH22 heliphorus
32	21	100.0	116	5	09NH29 miletus anc
33	21	100.0	117	5	09VEE2 drosophila
34	21	100.0	117	16	08YM59 anabaena sp
35	21	100.0	119	10	09FW06 arabidopsis
36	21	100.0	122	5	09NBL8 symbrenthia
37	21	100.0	123	11	09D9Z7 mus musculu
38	21	100.0	125	17	09TAY4 aeropyrum p
39	21	100.0	126	2	093UN4 acinetobact
40	21	100.0	126	5	095VL4 symbrenthia
41	21	100.0	127	17	096ZM5 sulfolobus
42	21	100.0	127	17	08TIB2 mechanosarc
43	21	100.0	130	11	070440 mus musculu
44	21	100.0	132	5	08WT32 leishmania
45	21	100.0	132	5	09BUC8 leishmania

ALIGNMENTS

RESULT 1
ID 09X142 PRELIMINARY; PRT; 31 AA.
AC 09X142;
DT 01-NOV-1999 (TREMUREL. 12, Created)
DT 01-NOV-1999 (TREMUREL. 12, Last sequence update)
DT 01-MAR-2002 (TREMUREL. 20, Last annotation update)
DE Hypothetical protein TM1316.
GN TM1316.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwyn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M., Stewart A.M., Cotton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA Heidelberg J., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.";
RL Nature 399:323-329 (1999).
DR EMBL; AE001786; MADS36390.1; -.
DR TIGR; TM1316; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 31 AA; 3178 MW; 1B03705C506B34BC CRC64;

Query Match 100.0%; Score 21; DB 16; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPN 4
DB 8 VVPN 11

DR Pfam; PF02221; E1_DerP2_DerF2; 1.
 DR SMART; SM00737; ML; 1.
 KW Allergen; Signal.
 FT SIGNAL 1 15
 FT CHARIN 16 141
 FT DISULFID 23 132
 FT DISULFID 36 41
 FT DISULFID 87 92
 FT CARBOHYD 103 103
 SQ SEQUENCE 141 AA; 14851 MW; 38EP9520010A04C1 CRC64; N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match Best/Local Similarity 100.0%; Score 21; DB 1; Length 141;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVPN 4
 111
 Db 111 VVPN 114

Search completed: September 4, 2003, 21:01:40
 Job time : 4.47059 secs

KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
 KM Multigene family.
 SQ SEQUENCE 132 AA; 13892 MW; B53FE2DA20970F43 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 132;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 ||||
 Db 109 VVPN 112

RESULT 13
 VP17_BPAPS
 ID VP17_BPAPS STANDARD; PRT; 138 AA.
 AC 09PTIT;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative protein p17.
 GN 17.
 OS Bacteriophage APSE-1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
 OX NCBI_TaxID=106199;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99420383; PubMed=10489345;
 RA van der Wilk F., Dullmanns A.M., Verbeek M., van den Heuvel J.F.J.M.;
 RT "Isolation and characterization of APSE-1, a bacteriophage infecting
 RT the secondary endosymbiont of acyrtosiphon pisum.";
 RL Virology 262:104-113(1999).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF157835; AAF03960.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 138 AA; 15063 MW; F0DE3B9E9F1FDE5B CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 138;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 ||||
 Db 35 VVPN 38

RESULT 14
 GSPG_KLEPN
 ID GSPG_KLEPN STANDARD; PRT; 140 AA.
 AC p15746;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE General secretion pathway protein G precursor (Pullulanase secretion
 DE protein p16).
 GN p16.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UNF 5023;
 RX MEDLINE=91109698; PubMed=2129543;
 RA Reys I., Pugsley A.P.;

RT "Five additional genes in the pulC-O operon of the Gram-negative
 RT bacterium *Klebsiella oxytoca* UNF5023 which are required for
 RT pullulanase secretion.";
 RL Mol. Gen. Genet. 222:176-184(1990).
 CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
 CC EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLLOCATION OF PULLULANASE.
 CC -1- SIMILARITY: BELONGS TO THE PULG/OUTG/PSG/XEG/XCP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M32613; AAA25129.1; -
 DR InterPro: IPR000983; Bac_GSPG.
 DR InterPro: IPR002416; Bac_GSPH.
 DR InterPro: IPR001120; Prok_N_methyltn.
 DR PRINTS: PRO0813; BCTERIALGSPG.
 DR PRINTS: PRO0885; BCTERIALGSPH.
 DR PROSITE: PS00409; PROKAR_NTER_METHYL, 1.
 KW Transport; Methylation.
 FT PROSEP 1 6 BY SIMILARITY.
 FT CHAIN 7 140 GENERAL SECRETION PATHWAY PROTEIN G.
 FT MOD_RES 7 7 METHYLATION (BY SIMILARITY).
 SQ SEQUENCE 140 AA; 15373 MW; 44264CC2B5A365B0 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 140;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 ||||
 Db 26 VVPN 29

RESULT 15
 ALL2_TYRPU
 ID ALL2_TYRPU STANDARD; PRT; 141 AA.
 AC 002380;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mite group 2 allergen Tyr p 2 precursor.
 OS Tyrophagus putrescentiae (Dust mite).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Acariformes; Sarcoplitiformes; Astigmata; Acaridea; Acaridae;
 OC Tyrophagus.
 OX NCBI_TaxID=59818;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-22.
 RX MEDLINE=98151280; PubMed=9492316;
 RA Eriksson T.L.J., Johansson E., Whitley P., Schmidt M., Eleyed S.,
 RA van Hage-Hamsten M.;
 RT "Cloning and characterization of a group II allergen from the dust
 RT mite *Tyrophagus putrescentiae*.";
 RL Eur. J. Biochem. 251:443-447(1998).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE NPC2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y12690; CAA73221.1; -
 DR HSP: Q00855; IAHK.
 DR InterPro: IPR003172; EL_DerP2_DerF2.

AC P04909;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A-alpha (H2A.1).
 GN H2A1 OR SPCC622.08C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NX NCBI_TaxID=4896;
 RX MEDLINE FROM N.A.
 RX MEDLINE=66310796; PubMed=3018512;
 RT "Organization, primary structure, and evolution of histone H2A and H2B genes of the fission yeast Schizosaccharomyces pombe."
 RL Mol. Cell. Biol. 5:3261-3269(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66135992; PubMed=4092687;
 RA Matsunoto S., Yanagida M.;
 RT "Histone gene organization of fission yeast: a common upstream sequence."
 RL EMO J. 4:3531-3538(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Brooks K., Brown D., Hayles J., Baker S., Basham D., Bowman S., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Butler S., Saunders D., Seeger K., Sharp S., Skellton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B., Weltjens I., Vanstreels E., Rieger M., Scheier K., Mueller-Auer S., Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M., Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Motier S., Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules each of H2A, H2B, H3 and H4. The octamer wraps approximately 146 bp of DNA.
 CC -1- SIMILARITY: Belongs to the histone H2A family.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: M1494; AAA35311.1;
 DR EMBL: X05220; CAA28848.1;
 DR EMBL: AL033127; CAA21864.1;
 DR PIR: B27399; H52PA2.

DR GeneDB_Spomb; SPCC622.08C;
 DR InterPro: IPR004822; Histone_core.
 DR InterPro: IPR002119; Histone_H2A.
 DR Pfam: PF00125; histone; 1.
 DR PRINTS: PR00620; HISTONEH2A.
 DR PRODOM: PD000522; Histone_H2A; 1.
 DR SMART: SM00414; H2A; 1.
 DR PROSITE: PS00046; HISTONE_H2A; 1.
 DR Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding; Multigene family.
 FT INIT MET 0
 FT CONFLICT 123 123 R -> G (IN REF. 1).
 SO SEQUENCE 131 AA; 13747 MW; DIAAF80F580C3273 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 131;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VVPN 4
 DB 108 VVPN 111
 RESULT 12
 H2A1 LEIDO STANDARD; PRT; 132 AA.
 ID H2A1 LEIDO
 AC P27891;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A.1.
 OS Leishmania donovani.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NX NCBI_TaxID=5661;
 RX MEDLINE=92209505; PubMed=1555581;
 RA Soto M., Requena J.M., Gomez L.C., Navarrete I., Alonso C.;
 RT "Molecular characterization of a Leishmania donovani infantum antigen identified as histone H2A."
 RL Eur. J. Biochem. 205:211-216(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Subsp. Infantum / LEM75 zymodeme 1;
 RX MEDLINE=92209505; PubMed=1555581;
 RA Soto M., Requena J.M., Jimenez-Rulz A., Alonso C.;
 RT "The mRNA coding for the nucleosomal protein H2A of Leishmania is polyadenylated and has stem-loops at the 3' end."
 RL Nucleic Acids Res. 19:4554-4554(1991).
 CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules each of H2A, H2B, H3 and H4. The octamer wraps approximately 146 bp of DNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the histone H2A family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: X60054; CAA42652.1;
 DR PIR: S22303; S22303.
 DR InterPro: IPR004822; Histone_core.
 DR InterPro: IPR002119; Histone_H2A.
 DR Pfam: PF00125; histone; 1.
 DR PRINTS: PR00620; HISTONEH2A.
 DR PRODOM: PD000522; Histone_H2A; 1.
 DR SMART: SM00414; H2A; 1.
 DR PROSITE: PS00046; HISTONE_H2A; 1.

```

RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shultz G., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Futerback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Flute L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.,
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT rd.";
RL Science 269:496-512(1995).
RN IDENTIFICATION.
RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D., Peterson J.,
RA Hickey E., Dodson R., Gwinn M.,
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO INSERTION ELEMENT IS1016 TRANSPOSASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32812; AAC22982.1; -.
DR TIGR: H11328.1; -.
KW Hypothetical protein; Transposon; Transposable element;
KM DNA recombination; Complete proteome.
SQ SEQUENCE 123 AA; 14466 MW; 7EBB8CD207AC622A CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
Db 9 VVPN 12

RESULT 10
H2A2_SCHPO STANDARD; PRT; 130 AA.
ID H2A2_SCHPO STANDARD; PRT; 130 AA.
AC P04910;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H2A-beta (H2A.2).
GN H2A2 OR SPAC19612.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN SEQUENCE FROM N.A.
RP MEDLINE=86310796; PubMed=3018512;
RA Choe J., Schuster T., Grunstein M.;
RT "Organization, primary structure, and evolution of histone H2A and
RT H2B genes of the fission yeast Schizosaccharomyces pombe.";
RL Mol. Cell. Biol. 5:3261-3269(1985).
RN SEQUENCE FROM N.A.
RP MEDLINE=86135992; PubMed=4092687;
RA Matsumoto S., Yanagida M.;
RT "Histone gene organization of fission yeast: a common upstream
RT sequence.";
RL EMBL J. 4:3531-3538(1985).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream W.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gantley S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Hunkle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickert G., Aert R., Rodden J., Gymnopoulos B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules
CC each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC bp of DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the histone H2A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M11500; AAA35310.1; -.
DR EMBL: X05221; CA28849.1; -.
DR EMBL: Z97209; CAB10117.1; -.
DR PTR: C27399; HSZPA3.
DR GenBank_Spomb; SPAC19612.06C; -.
DR InterPro: IPR004822; Histone_core.
DR InterPro: IPR002119; Histone_H2A.
DR Pfam: PF00125; histone.1.
DR PRINTS: PR00620; HISTONEH2A.
DR ProDom: PD000522; Histone_H2A.1.
DR SMART: SM00414; H2A; 1.
DR PROSITE: PS00046; HISTONE_H2A; 1.
KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW Multigene family.
FT INIT_MET 0
FT SEQUENCE 130 AA; 13645 MW; 1BD5C8B5C8692A04 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 130;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
Db 108 VVPN 111

RESULT 11
H2A1_SCHPO STANDARD; PRT; 131 AA.
ID H2A1_SCHPO STANDARD; PRT; 131 AA.

```

```

DE Protein B20.
GN B20R.
OS Vaccinia virus (strain WR).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91259063; PubMed=2045793;
RA Smith G.L., Chan Y.S., Howard S.T.;
RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RL the right inverted terminal repeat.";
RL J. Gen. Virol. 72:1349-1376(1991).
CC -1 SIMILARITY: Contains 1 ANK repeat.
CC -1 CAUTION: B20 IN WR DOES NOT CORRESPOND TO B20 IN COPENHAGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D11079; BAA01850.1;
DR PIR: J01814; J01814.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank. 1.
DR SMART: SM00248; ANK. 1.
DR PROSITE: PS50088; ANK_REPEAT: 1.
DR PROSITE: PS50297; ANK_REGION: 1.
KW ANK repeat.
FT REPEAT 75 104 ANK.
SQ SEQUENCE 115 AA: 13439 MW: BAD72B66F1335F1A CRC64:

Query Match 100.0%; Score 21; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPN 4
DB 103 VVPN 106

RESULT 8
MAR1_HUMAN STANDARD: PRT; 118 AA.
AC Q16653;
DR 01-NOV-1997 (Rel. 35, Created)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Melanoma antigen recognized by T-cells 1 (MART-1) (Melan-A protein)
DE (Antigen SK29-AA) (Antigen LB39-AA).
GN MLANA OR MART1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=94224770; PubMed=8170938;
RA Kawakami Y., Elyahu S., Delgado C.H., Robbins P.F., Rivoltini L.,
RA Topalian S.L., Miki T., Rosenberg S.A.;
RT "Cloning of the gene coding for a shared human melanoma antigen
RT recognized by autologous T cells infiltrating into tumor.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3515-3519(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94275389; PubMed=8006593;
RA Coule P.G., Brichard V., van Pel A., Moelfel T., Schneider J.,
RA Traversari C., Mattel S., de Plaen E., Lurquin C., Szikora J.-P.,
RA Renaud J.-C., Boon T.;

```

```

RT "A new gene coding for a differentiation antigen recognized by
RT autologous cytolytic T lymphocytes on HLA-A2 melanomas.";
RL J. Exp. Med. 180:35-42(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Mausina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bawa S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maitra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSION IS RESTRICTED TO MELANOMA AND
CC MELANOCYTE CELL LINES AND RETINA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U06452; AAA19238.1;
DR EMBL: U06654; AAA20389.1;
DR EMBL: BC014423; AAH14423.1;
DR PIR: A55253; A55253.
DR GeneW: HGNC:7124; MLANA.
DR MIM: 605513;
DR GO: GO:0005887; C: integral to plasma membrane; TAS.
DR GO: GO:0008222; F: tumor antigen; TAS.
KW Antigen; Transmembrane.
FT TRANSMEM 27 47 POTENTIAL.
SQ SEQUENCE 118 AA: 13157 MW: B755BF39CFB16E CRC64:

Query Match 100.0%; Score 21; DB 1; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVPN 4
DB 96 VVPN 99

RESULT 9
YD2A_HAETN STANDARD: PRT; 123 AA.
AC O86236;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transposase-like protein H11328.1.
GN H11328.1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]

```

```

ID      HFO_WIGBR      STANDARD:      PRT;      87 AA.
AC      Q8D317;
DT      15-SEP-2003 (Rel. 42, Created)
DT      15-SEP-2003 (Rel. 42, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Hfq protein.
GN      HFO OR WIGBR1840.
OS      Wiglesworthia glossinidia brevipalpis.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Wiglesworthia.
OX      NCBI_TaxID=36870;
RN
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22297718; PubMed=12219091;
RA      Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA      Aksoy S.;
RT      "Genome sequence of the endocellular obligate symbiont of tsetse
RT      files, Wiglesworthia glossinidia.";
RL      Nat. Genet. 32:402-407(2002).
CC      -!- FUNCTION: RNA-binding protein that stimulates the elongation of
CC      poly(A) tails (By similarity).
CC      -!- SIMILARITY: BELONGS TO THE HFO FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; AB063521; BAC24330.1; -
DR      HAMAP; MF_00436; -; 1.
DR      InterPro: IPR005001; Hfq.
DR      Pfam: PF03329; Hfq; 1.
KW      RNA-binding, Complete, proteome.
SQ      SEQUENCE 87 AA; 9975 MW; F9239AF6EC42DB2 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 VVPN 4
        1111
Db      62 VVPN 65

RESULT 6
THIO_MYCTU      STANDARD:      PRT;      115 AA.
ID      P52229;
AC      P52229;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Thioredoxin (TRX) (MPT46).
GN      TRXA OR TRX OR TRXC OR RV3914 OR MT4033 OR MTW028.05.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Wiles B., Phillip W., Drifhout J.W., Offringa R.,
RA      Ottenhoff T.H.M.;
RL      Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,
RA      Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies K., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,

```

```

RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrall B.G.,
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence.";
RL      Nature 393:537-544(1998).
[3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.;
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains.";
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[4]
RP      SEQUENCE OF 1-20.
RX      MEDLINE=96071926; PubMed=7591163;
RA      Wiles B., Nagai S., Wiker H.G., Harboe M., Ottenhoff T.H.M.;
RT      "Identification and functional characterization of thioredoxin of
RT      Mycobacterium tuberculosis.";
RL      Infect. Immun. 63:4946-4948(1995).
CC      -!- FUNCTION: Participates in various redox reactions through the
CC      reversible oxidation of its active center dithiol to a disulfide
CC      and catalyzes dithiol-disulfide exchange reactions (By
CC      similarity).
CC      -!- SIMILARITY: BELONGS TO THE THIOREDIXIN FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; X95798; CA65071.1; -
DR      EMBL; AL021426; CA16227.1; -
DR      EMBL; AE007194; AAK48398.1; -
DR      PIR; B70851; B70851.
DR      HSSP; P80579; 1QWV.
DR      TIGR; MT4033; -.
DR      Tuberculist; RV3914; -
DR      InterPro: IPR006662; Thiored.
DR      InterPro: IPR006663; Thioredox_dom2.
DR      InterPro: IPR005746; Thioredoxin.
DR      Pfam; PF00085; thiored; 1.
DR      TIGRFAMS; TIGR01068; thioredoxin; 1.
KW      PROSITE; PS00194; THIOREDIXIN; 1.
KW      Redox-active center; Electron transport; Complete proteome.
FT      INIT_MET 0
FT      DISULFID 36 39
SQ      SEQUENCE 115 AA; 12413 MW; 9B369AFD2FARD8C7 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 VVPN 4
        1111
Db      110 VVPN 113

RESULT 7
VB20_VACCV
ID      VB20_VACCV      STANDARD:      PRT;      115 AA.
AC      Q01224;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      01-OCT-2001 (Rel. 40, Last annotation update)

```


RT of the protein complexes of plant mitochondria."
 RL Plant J. 9:357-368(1996).
 CC -1- FUNCTION: THIS PROTEIN SEEMS TO BE PART OF THE STALK THAT LINKS
 CC CF(0) TO CF(1). IT EITHER TRANSMITS CONFORMATIONAL CHANGES FROM
 CC CF(0) INTO CF(1) OR IS IMPLICATED IN PROTON CONDUCTION.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(In) -> ADP + phosphate +
 CC H(+)(Out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.
 DR InterPro: IPR000711; ATPsynth_OSCP.
 DR PROSITE: PS00389; ATPase_Delta; PARTIAL.
 KM Hydrolyase; ATP synthetase; CF(1); Hydrogen ion transport;
 KM Mitochondrion.
 FT NON TER 37
 SQ SEQUENCE 37 AA: 4003 MW: 98FDAB1A4298F4AF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 37;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 1111
 DB 21 VVPN 24

RESULT 3

NUFM_SOLUTU STANDARD: PRT; 37 AA.
 AC P80266;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-ubiquinone oxidoreductase 22.5 kDa subunit (EC 1.6.5.3)
 DE (EC 1.6.99.3) (Complex I-22.5kD) (CI-22.5kD) (Fragment).
 OS Solanum tuberosum (Potato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. Bintje; TISSUE=tuber;
 RX MEDLINE=94124587; PubMed=8294484;
 RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
 RA Grohmann L.,
 RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
 RT the respiratory chain from the inner mitochondrial membrane of
 RT Solanum tuberosum."
 RL J. Biol. Chem. 269:2263-2269(1994).
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE. THIS IS A COMPONENT OF THE IRON-SULFUR (IR)
 CC FRAGMENT OF THE ENZYME.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I NDUFAS SUBUNIT FAMILY.
 DR PIR: F49732; F49732.
 DR InterPro: IPR006806; ETC_CI_29_9.
 DR Pfam: PF04716; ETC_CI_29_9; 1.
 KM Oxidoreductase; NAD: Ubiquinone; Mitochondrion.
 FT NON TER 37
 SQ SEQUENCE 37 AA: 4073 MW: FB5C5A7897C60928 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 37;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 1111
 DB 14 VVPN 17

RESULT 4

AT19_YEAST STANDARD: PRT; 68 AA.
 AC P81451;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP synthase k chain, mitochondrial (EC 3.6.3.14).
 GN ATP19 OR YOL078W OR YOL077W-A.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=w303-1A.
 RX MEDLINE=97321807; PubMed=9178509;
 RA Tzermia M., Katsoulou C., Alexandraki D.;
 RA Arnold I., Pfeiffer K., Neupert W., Stuart R.A., Schaegeer H.;
 RT "Sequence analysis of a 33.2 kb segment from the left arm of yeast
 RT chromosome XV reveals eight known genes and ten new open reading
 RT frames including homologues of ABC transporters, inositol
 RT phosphatases and human expressed sequence tags."
 RL Yeast 13:583-589(1997).
 RN [2]
 RP IDENTIFICATION, AND SEQUENCE OF 1-14.

RC STRAIN=w303-1A.
 RX MEDLINE=99077785; PubMed=9857174;
 RA Arnold I., Pfeiffer K., Neupert W., Stuart R.A., Schaegeer H.;
 RT "Yeast mitochondrial F1F0-ATPase exists as a dimer: identification of
 RT three dimer-specific subunits."
 RL EMBO J. 17:7170-7178(1998).
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX. THE K CHAIN
 CC BINDS THE DIMERIC FORM BY INTERACTING WITH THE G AND E CHAINS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(In) -> ADP + phosphate +
 CC H(+)(Out).
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. IN YEAST, THE
 CC DIMERIC FORM OF ATP SYNTHASE CONSISTS OF 18 POLYPEPTIDES: ALPHA,
 CC BETA, GAMMA, DELTA, EPSILON, 4 (B), 5 (OSCP), 6 (A), 8, 9 (C), D,
 CC E (TIM11), F, G, H, I, J AND K.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Z74820; -; NOT_ANNOTATED_CDS.
 DR PIR: S78739; S78739.
 DR SGD: S0007339; ATP19.
 DR GO: GO:0000276; C:proton-transporting ATP synthase complex, c...; IMP.
 DR GO: GO:0015986; P:ATP synthetase coupled proton transport; IMP.
 KM ATP synthetase; Hydrogen ion transport; Mitochondrion; CF(0).
 SQ SEQUENCE 68 AA: 7534 MW: 4FE05C1D93501283 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 68;
 Best Local Similarity 100.0%; Pred. No. 118+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVPN 4
 1111
 DB 29 VVPN 32

RESULT 5

HFO_MIGBR

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 : Search time 3.47059 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-55
Perfect score: 21
Sequence: 1 VVPN 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt.41:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	20	1	OMPW_VIBAL
2	21	100.0	37	1	ATPO_SOLTU
3	21	100.0	37	1	NUFM_SOLTU
4	21	100.0	68	1	AT19_YEAST
5	21	100.0	87	1	HEQ_WIGBR
6	21	100.0	115	1	THIO_MXCTU
7	21	100.0	115	1	VB20_VACCV
8	21	100.0	118	1	MAR1_HUMAN
9	21	100.0	123	1	YD2A_HAEN
10	21	100.0	130	1	H2A2_SCHPO
11	21	100.0	131	1	H2A1_SCHPO
12	21	100.0	132	1	H2A1_LEIDO
13	21	100.0	138	1	VP17_BPAPS
14	21	100.0	140	1	GSPG_KLEPN
15	21	100.0	141	1	AL12_TYRPU
16	21	100.0	143	1	GSPG_AERHY
17	21	100.0	144	1	URE2_YERPE
18	21	100.0	145	1	GSPG_ECOLI
19	21	100.0	145	1	NINB_BP22
20	21	100.0	146	1	GSPG_VIBCH
21	21	100.0	146	1	NINB_ECOL6
22	21	100.0	146	1	NINB_LAMBD
23	21	100.0	146	1	VA28_VACCV
24	21	100.0	148	1	R19G_ASCSU
25	21	100.0	153	1	GSPG_ERWCH
26	21	100.0	155	1	IIF5_HUMAN
27	21	100.0	156	1	GSPG_ERWCA
28	21	100.0	156	1	IIF5_MOUSE
29	21	100.0	160	1	KAFK_PHRPO
30	21	100.0	160	1	PETD_CHLRE
31	21	100.0	160	1	PETD_CHLVD
32	21	100.0	160	1	PETD_PROHO
33	21	100.0	162	1	Y4UG_RHISN

34	21	100.0	163	1	URE2_YEREN	P31495 yersinia en
35	21	100.0	169	1	NUFM_ARATH	G9FLX7 arabidopsis
36	21	100.0	172	1	ESP4_LACVY	P35578 lacerta viv
37	21	100.0	176	1	CYTL_STOHE	P81662 stochactis
38	21	100.0	180	1	SP22_CANFA	P12280 canis fam1
39	21	100.0	180	1	SP22_CHICK	P28687 gallus gall
40	21	100.0	180	1	SP22_HUMAN	G9H057 homo sapien
41	21	100.0	185	1	RM05_BRANA	P49388 brassica na
42	21	100.0	187	1	Y418_VIBCH	G9KUN7 vibrio chol
43	21	100.0	189	1	YWOC_BACSU	P94573 bacillus su
44	21	100.0	195	1	KITH_BACSU	G03221 bacillus su
45	21	100.0	196	1	NODC_RHISM	P04679 rhizobium s

ALIGNMENTS

RESULT 1
OMPW_VIBAL STANDARD: PRT; 20 AA.
ID P83151:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Outer membrane protein W (Outer membrane protein 25Va) (Omp25Va)
DE (Fragment).
GN OMPW
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=663;
RN [1]
RP SEQUENCE.
RC STRAIN=NCIMB 19037;
RA Onji M., Hirabayashi J., Suzuki S.;
RT "Characterization of major outer membrane proteins of Vibrio
RT alginolyticus and the stability against proteases.";
RL Microbes Environ. 0:0-0(2002).
CC -1- SUBCELLULAR LOCATION: Outer membrane.
CC -1- SIMILARITY: BELONGS TO THE OMPW/ALKL FAMILY.
KW Outer membrane.
FT NON_TER
SQ SEQUENCE 20 AA; 2096 MW; D29PE7FCA16CD37 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VVPN 4
Db 15 VVPN 18
RESULT 2
ATPO_SOLTU STANDARD: PRT; 37 AA.
ID P80504:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase delta chain, mitochondrial (EC 3.6.3.14) (Oligomycin
DE sensitivity conferral protein) (OSCP) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC TISSUE= tuber;
RC MEDLINE=97077345; PubMed=8919912;
RA Jansch L., Kruft V., Schmitz U.K., Braun H.P.;
RT "New insights into the composition, molecular mass and stoichiometry

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 9.35294 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893d-57

Perfect score: 31

Sequence: 1 GAVVPN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	468	2 T22397	hypothetical prote
2	31	100.0	1360	2 T12064	DNA binding protei
3	30	96.8	72	2 T12130	potassium channel
4	30	96.8	72	2 T03391	potassium channel
5	30	96.8	99	1 W5BP57	gene 5.5 protein -
6	30	96.8	340	2 F95266	probable lact-fam1
7	30	96.8	587	1 W2BEC9	gene 36 protein -
8	30	96.8	587	2 T42580	gene 36 protein -
9	30	96.8	803	1 MNV94	early 94K protein
10	30	96.8	803	2 G72866	hypothetical prote
11	30	96.8	807	2 T12177	potassium channel
12	30	96.8	1891	2 T43262	calcium channel al
13	28	90.3	190	2 E84355	DNA-directed RNA p
14	28	90.3	339	2 A13609	daunorubicin resis
15	28	90.3	425	2 A87631	FMN oxidoreductase
16	28	90.3	425	2 T18335	hypothetical prote
17	28	90.3	428	2 C26532	5-enolpyruvylshik
18	28	90.3	468	2 C95405	probable sulfite o
19	28	90.3	469	2 T46929	hypothetical prote
20	28	90.3	469	2 T46930	hypothetical prote
21	28	90.3	530	2 S52215	hypothetical prote
22	28	90.3	547	2 T06758	probable galactose
23	28	90.3	569	1 A24984	ribulokinase (Im
24	28	90.3	569	2 AD0515	L-ribulokinase (Im
25	28	90.3	586	2 T08293	hypothetical prote
26	28	90.3	636	2 D84823	hypothetical prote
27	27	87.1	66	2 S03595	dhfr protein - phag
28	27	87.1	108	2 B43936	ORF 3' of clsa - B
29	27	87.1	131	1 H52PA3	histone H2A.2 - fl

30	27	87.1	132	1 H52PA2	histone H2A.1 - fl
31	27	87.1	213	2 D71669	adenylylate kinase (
32	27	87.1	239	2 F81704	serine esterase, p
33	27	87.1	239	2 A71552	probable lysophosp
34	27	87.1	359	2 F86683	prophage pil prote
35	27	87.1	451	2 T32777	hypothetical prote
36	27	87.1	483	2 F64760	membrane protein p
37	27	87.1	483	2 C90677	hypothetical prote
38	27	87.1	483	2 F85527	hypothetical prote
39	27	87.1	483	2 A10547	PIPD protein (limp
40	27	87.1	537	2 H90608	conserved hypothet
41	27	87.1	620	2 T50232	actin-like protein
42	27	87.1	682	2 T12968	hypothetical prote
43	27	87.1	711	2 T30107	hypothetical prote
44	27	87.1	744	2 T45943	hypothetical prote
45	27	87.1	767	1 S50594	5-methyltetrahydro

ALIGNMENTS

RESULT 1

T22397 hypothetical protein F49A5.7 - Caenorhabditis elegans

C.Species: Caenorhabditis elegans

C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C.Accession: T22397

R.Motifmore, B.

submitted to the EMBL Data Library, November 1996

A.Reference number: Z19560

A.Accession: T22397

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-468 <M11>

A.Cross-References: EMBL:Z81542; PIDN:CA804419.1; GSPDB:GN00023; CESP:F49A5.7

A.Experimental source: clone F49A5

A.Genetics:

A.Gene: CESP:F49A5.7

A.Map position: 5

A.Intons: 19/1; 66/1; 151/1; 174/1; 189/1; 224/3; 248/3; 297/1; 340/1

Query Match 100.0%; Score 31; DB 2; Length 468;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6

DB 401 GAVVPN 406

RESULT 2

T12064 DNA binding protein G11-1 - African clawed frog (fragment)

C.Species: Xenopus laevis (African clawed frog)

C.Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Nov-2000

C.Accession: T12064

R.Ruitz i Altaba, A.; Lee, J.D.

submitted to the EMBL Data Library, May 1996

A.Reference number: Z17399

A.Accession: T12064

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-1360 <RUI>

A.Cross-References: EMBL:U57454; NID:q3282202; PID:q3282203

C:Superfamily: g11 transforming protein

C:Keywords: zinc finger

Query Match 100.0%; Score 31; DB 2; Length 1360;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6

DB 401 GAVVPN 406

Db 1332 GAVVFN 1337

RESULT 3

T12130

potassium channel protein - fava bean (fragment)

C:Species: Vicia faba (fava bean)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Dec-2002

C:Accession: T12130

R:Phillippar, K.

submitted to the EMBL Data Library, December 1996

A:Reference number: Z17429

A:Accession: T12130

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-72 <PHI>

A:Cross-references: EMBL:Y09749; NID:e1012614; PID:e293987

A:Experimental source: leaf

C:Superfamily: Arabidopsis potassium channel protein AKT1; ankyrin repeat homology; CAMH

C:Keywords: potassium channel; potassium transport; transmembrane protein; transport pro

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 72;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y

1 GAVVFN 6

|||||

Db 5 GAVVFN 10

RESULT 4

T03391

potassium channel - maize (fragment)

C:Species: Zea mays (maize)

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000

C:Accession: T03391

R:Holth, S.; Dreyer, I.; Dietrich, P.; Becker, D.; Mueller-Roeber, B.; Hedrich, R.

Proc. Natl. Acad. Sci. U.S.A. 94, 4806-4810, 1997

A:Title: Molecular basis of plant-specific acid activation of K⁺ uptake channels.

A:Reference number: Z14922; MUID:97272307; PMID:9114073

A:Accession: T03391

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-72 <HO>

A:Cross-references: EMBL:Y09747; NID:g2104905; PIDN:CAA70894.1; PID:g2104906

A:Experimental source: strain I5511B; coleoptile

C:Keywords: potassium channel; potassium transport; transmembrane protein

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 72;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y

1 GAVVFN 6

|||||

Db 5 GAVVFN 10

RESULT 5

W5B957

gene 5.5 protein - phage T7

C:Species: Phage T7

C>Date: 13-Jun-1983 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000

C:Accession: A04412; S42313

R:Dunn, J.J.; Thompson, K.

submitted to the Nucleic Acid Sequence Database, September 1982

A:Reference number: A94615

A:Accession: A04412

A:Molecule type: DNA

A:Residues: 1-99 <DUN>

R:Dunn, J.J.; Studier, F.W.

J. Mol. Biol. 166, 477-535, 1983

A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge

A:Reference number: S42283; MUID:83241725; PMID:6864790

A:Accession: S42313

A:Molecule type: DNA

A:Residues: 1-99 <DUN>

A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24414.1; PID:g15593

A>Note: the authors did not translate the codon for residue 1

C:Genetics:

A:Gene: 5.5

A:Map position: 42,20-43,46

C:Superfamily: phage T7 gene 5.5 protein

Query Match

Best Local Similarity 96.8%; Score 30; DB 1; Length 99;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y

1 GAVVFN 6

|||||

Db 39 GAVVFN 44

RESULT 6

F95266

probable lacI-family transcription regulator [Imported] - Sinorhizobium meliloti (str

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: F95266

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B

.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: F95266

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK64696.1; PID:g14523096; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSYMA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Punter, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pel, D.; Chaim, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau

hebaul, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma0078

A:Genome: plasmid

C:Superfamily: lac repressor

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 340;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y

1 GAVVFN 6

|||||

Db 72 GAVVFN 77

RESULT 7

WZBB9

gene 36 protein - equine herpesvirus 1 (strain Ab4p)

C:Species: equine herpesvirus 1

A>Note: host Equus caballus (domestic horse)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: B36799

R:Gelford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to Genbank, March 1992

A:Description: The DNA sequence of equine herpesvirus-1.

A:Reference number: A36805

A:Accession: B36799

A:Molecule type: DNA

A:Residues: 1-587 <TEL>

A:Cross-references: GB:M8664; NID:9330791; PIDN:AA02472.1; PID:9330829
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1
A:Reference number: M41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 36
C:Superfamily: varicella-zoster virus gene 34 protein

Query Match 96.8%; Score 30; DB 1; Length 587;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
||:||||
Db 413 GAVPN 418

RESULT 8
T42580
gene 36 protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42580
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: Z22173; MUID:98264497; PMID:9603335
A:Accession: T42580
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-587 <TEL>
A:Cross-references: EMBL:AF030027; NID:92605950; PIDN:AAC59553.1; PID:92605981
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 36
C:Superfamily: varicella-zoster virus gene 34 protein

Query Match 96.8%; Score 30; DB 2; Length 587;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
||:||||
Db 413 GAVPN 418

RESULT 9
NMNV94
early 94K protein - Autographa californica nuclear polyhedrosis virus (strain L-1)
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: B27840
R:Friesen, P.D.; Miller, L.K.
J. Virol. 61, 2264-2272, 1987
A:Title: Divergent transcription of early 35- and 94-kilodalton protein genes encoded by
A:Reference number: A93026; MUID:87226411; PMID:3035225
A:Accession: B27840
A:Molecule type: DNA
A:Residues: 1-803 <FRI>
A:Cross-references: GB:M16821; NID:9332445; PIDN:AAA6702.1; PID:9332446
C:Superfamily: Autographa californica nuclear polyhedrosis virus early 94K protein
C:Keywords: early protein

Query Match 96.8%; Score 30; DB 1; Length 803;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
||:||||

Db 412 GAVPN 417

RESULT 10
G72866
hypothetical protein - Autographa californica nuclear polyhedrosis virus.
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-Jan-2000
C:Accession: G72866
R:AYres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: G72866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <AYR>
A:Cross-references: GB:L22858; NID:9510708; PIDN:AAA66764.1; PID:9559203
A:Gene: Ac-94K
C:Superfamily: Autographa californica nuclear polyhedrosis virus early 94K protein

Query Match 96.8%; Score 30; DB 2; Length 803;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
||:||||
Db 412 GAVPN 417

RESULT 11
T12177
potassium channel protein - fava bean (fragment)
C:Species: Vicia faba (fava bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Dec-2002
C:Accession: T12177
R:Arche, P.; Wohlfarth, T.; Hedrich, R.; Becker, D.
submitted to the EMBL Data Library, January 1997
A:Reference number: Z17443
A:Accession: T12177
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-807 <ACH>
A:Cross-references: EMBL:Y10579; NID:e1015303; PID:9321441
A:Experimental source: cotyledon; clone VKC1
C:Superfamily: Arabidopsis potassium channel protein AKT1; ankyrin repeat homology; c
C:Keywords: potassium channel; transmembrane protein; transport protein; voltage-gate
F:550-582/Domain: ankyrin repeat homology <ANK>

Query Match 96.8%; Score 30; DB 2; Length 807;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
||:||||
Db 230 GAVPN 235

RESULT 12
T43262
calcium channel alpha-1 chain, L-type - Stylophora pistillata
C:Species: Stylophora pistillata
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T43262
R:Zoccola, D.; Tambutte, E.; Senegas-Bajas, F.; Michiels, J.F.; Falla, J.P.; Jaubert
Gene 227, 157-167, 1999
A:Title: Cloning of a calcium channel alpha1 subunit from the reef-building coral, St
A:Reference number: Z22375; MUID:99148007; PMID:10023047
A:Accession: T43262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-1891 <ZOC>
 A:Cross-references: EMBL:U64465; NID:g4204977; PID:g4204978; PIDN:AD11470.1
 C:Genetics:
 A:Gene: CACHL
 C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 96.8%; Score 30; DB 2; Length 1891;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 |||:|
 Db 744 GAVVFN 749

RESULT 13

E84355

DNA-directed RNA polymerase subunit E' [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: E84355

R:NM, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jadic, Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016505

A:Accession: E84355

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-190 <STO>

A:Cross-references: GB:AE004437; NID:g10581476; PIDN:AAG20209.1; GSPDB:GN00138

C:Genetics:

A:Gene: rpoE'

C:Superfamily: DNA-directed RNA polymerase subunit E

Query Match 90.3%; Score 28; DB 2; Length 190;
 Best Local Similarity 83.3%; Pred. No. 59;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 |||:|
 Db 57 GAVVFN 62

RESULT 14

A13609

daunorubicin resistance ATP-binding protein drrA BME110802 [imported] - Brucella melitens

C:Species: Brucella melitens

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: A13609

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3252; PMID:11756688

A:Accession: A13609

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-339 <KUR>

A:Cross-references: GB:AE008918; PIDN:AL54044.1; PID:g17984998; GSPDB:GN00191

C:Genetics:

A:Gene: BME110802

A:Map position: 11

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 90.3%; Score 28; DB 2; Length 339;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6

Db |||:|
 31 GAVVFN 36

RESULT 15

A87631

FMN oxidoreductase CC3083 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: A87631

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hatt, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87631

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-425 <STO>

A:Cross-references: GB:AE005673; NID:g13424735; PIDN:AAK25045.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3083

Query Match 90.3%; Score 28; DB 2; Length 425;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 |||:|
 Db 18 GAVVFN 23

Search completed: September 4, 2003, 21:10:48
 Job time : 10.3529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11; Search time 5.20588 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-57
Perfect score: 31
Sequence: 1 GAVPN 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt-41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	1360	1 GLI1_XENLA	Q91690 xenopus lae
2	30	96.8	169	1 V55_BPT7	P03787 bacterioph
3	30	96.8	587	1 UL25_HSVB	P28928 equine herp
4	30	96.8	803	1 VE94_NPVAC	P08161 autographa
5	28	90.3	377	1 PEB1_DROME	Q9u615 drosophila
6	28	90.3	428	1 AROA_BACSU	P20691 bacillus su
7	28	90.3	568	1 ARAB_SALTY	P58542 salmonella
8	28	90.3	568	1 ARAB_SALTY	P06188 salmonella
9	27	87.1	66	1 VDRH_BP186	P21680 bacterioph
10	27	87.1	130	1 H2A2_SCHPO	P04910 schizosacch
11	27	87.1	131	1 H2A1_SCHPO	P04909 schizosacch
12	27	87.1	153	1 ML1A_XENLA	P51048 xenopus lae
13	27	87.1	213	1 KAD_RICPR	Q9zc66 ticketulia
14	27	87.1	482	1 PRPD_ECOS7	Q8x693 escherichia
15	27	87.1	482	1 PRPD_ECOLI	P77243 escherichia
16	27	87.1	483	1 PRPD_SALTY	Q8z903 salmonella
17	27	87.1	483	1 PRPD_SALTY	P74840 salmonella
18	27	87.1	483	1 ARPB_SCHPO	Q9u607 schizosacch
19	27	87.1	766	1 METE_YEAST	P05694 saccharomyc
20	27	87.1	813	1 TTKA_DROME	P42282 drosophila
21	27	87.1	918	1 DNLI_RAT	Q91hy8 rattus norv
22	27	87.1	3063	1 POLG_PYYN	P18247 p genome po
23	26	83.9	145	1 H2A1_WHEAT	P02275 triticum ae
24	26	83.9	151	1 H2A2_WHEAT	P02276 triticum ae
25	26	83.9	151	1 H2A3_WHEAT	P02277 triticum ae
26	26	83.9	189	1 YWOC_BACSU	P94573 bacillus su
27	26	83.9	213	1 RL3_BACST	P28600 bacillus su
28	26	83.9	268	1 Y237_THEAC	Q9h1j1 thermoplasma
29	26	83.9	270	1 RM21_URATH	Q81940 arabidopsis
30	26	83.9	282	1 DRN1_HUMAN	P24855 homo sapien
31	26	83.9	287	1 M48A_MOUSE	Q99n10 mus musculu
32	26	83.9	292	1 E434_ADE09	P89083 human adeno
33	26	83.9	299	1 ADD_TREPA	O83085 treponema p

34	26	83.9	320	1 CHIX_PEA	P36907 pisum sativ
35	26	83.9	333	1 CCPA_STRMU	O07329 streptococc
36	26	83.9	336	1 CH12_ORYZA	P25765 oryza sativ
37	26	83.9	345	1 HAIR_CHICK	P15979 gallus gall
38	26	83.9	353	1 DPF1_HUMAN	Q92782 homo sapien
39	26	83.9	353	1 PSBA_PROHO	P15191 prochloroth
40	26	83.9	356	1 PSBA_CYAA5	P51759 cyanothecce
41	26	83.9	356	1 REOC_MOUSE	P58269 mus musculu
42	26	83.9	360	1 PSB1_ANASP	P46242 anabaena sp
43	26	83.9	360	1 PSB1_SYNEL	P35876 synecchococc
44	26	83.9	360	1 PSB1_SYNP7	P04996 synecchococc
45	26	83.9	360	1 PSB1_SYNVU	P51765 synecchococc

ALIGNMENTS

```

RESULT 1
GLI1_XENLA          STANDARD;          PRT; 1360 AA.
ID GLI1_XENLA
AC Q91690;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein GLI1 (GLI-1) (Fragment).
GN GLI1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97359568; PubMed=9216996;
RA Lee J., Platt K.A., Censullo P., Ruiz i Altaba A.;
RT "Gli1 is a target of Sonic hedgehog that induces ventral neural tube
RT development.";
RL Development 124:2537-2552(1997).
RN [2]
RP REVISIONS.
RA Lee J., Platt K.A., Censullo P., Ruiz i Altaba A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IMPLICATED IN THE TRANSDUCTION OF SHH SIGNAL (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U57454; AAC24946.1; -
DR F01R; T12064; T12064.
DR HSSP; P08151; ZGLI.
DR InterPro: IPR007087; znf_C2H2.
DR Pfam: PF00096; zf_C2H2; 5.
DR SMART: SM00355; znf_C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 5.
KW Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;
KW Nuclear protein; Repeat.
FT ZN_FING 250 275 C2H2-TYPE.
FT ZN_FING 316 340 C2H2-TYPE.
FT ZN_FING 346 371 C2H2-TYPE.
FT ZN_FING 377 402 C2H2-TYPE.
FT NON_TER 1360 1360
SQ
SDOUCNCE 1360 AA; 149422 MW; 5A32B8086794EC2D CRC64;

```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M16821; AAAA6702.1; -

DR EMBL: L22858; AAAA6764.1; -

DR PIR: B27840; WMNV94.

DR PIR: G72866; G72866.

KW Early protein.

FT CONFLICT 457 457 E -> K (TN REF. 1)

SQ SEQUENCE 803 AA; 94540 MM; A89964321382097D CRC64;

Query Match 96.8%; Score 30; DB 1; Length 803;

Best Local Similarity 83.3%; Pred. No. 42;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPM 6

DB 412 GAVVPM 417

RESULT 5

PEBI_DROME STANDARD; PRT; 377 AA.

ID PEBI_DROME

AC Q9U6L5; Q9U6M1;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ejaculatory bulb specific protein I precursor (PEB-me).

GN PEB OR BCDNA:GH06048 OR CG268.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RX MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gockayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.D., Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P., Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Egholm E., Fennell T., Ferraz C., Ferrelle S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwae C., Jalili M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Mishina N.V., Modarres C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy N., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reuter K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spredling A.C., Stapleton M., Strong R., Sun E., Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";

RL Science 287:2185-2195(2000).

RN [2]

RP REVISIONS.

RC STRAIN-Berkeley;

RX MEDLINE-22426059; PubMed-12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Belencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;

RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a RT systematic review.";

RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley; TISSUE-Head;

RX MEDLINE-20196012; PubMed-10731138;

RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E., Stapleton M., Harvey D.A.;

RT "A *Drosophila* complementary DNA resource.";

RL Science 287:2222-2224(2000).

RN [4]

RP SEQUENCE OF 77-90 AND 101-109, FUNCTION, SUBCELLULAR LOCATION, TISSUE RP SPECIFICITY, AND VARIANT.

RC STRAIN-Canton-S;

RX MEDLINE-21167515; PubMed-11267893;

RA Lung O., Wolfner M.F.;

RT "Identification and characterization of the major *Drosophila* RT melanogaster mating plug protein.";

RL Insect Biochem. Mol. Biol. 31:543-551(2001).

CC -1- FUNCTION: Major protein component of the posterior mating plug. CC Accessory gland proteins constitute, or are required for formation CC of the anterior mating plug. Posterior mating plug forms before CC sperm transfer and the anterior mating plug is formed after the CC start of mating.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Specifically expressed in the ejaculatory bulb CC and seminal fluid. Detected in mated females 3 minutes after the CC start of mating, and for at least 3 hours after the start of CC mating.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE003466; AAF47319.2; -

DR EMBL: AF184225; AAD55736.1; -

DR FlyBase: FBgn0004181; Pdb.

KW Behavior: Signal; Polymorphism.

FT SIGNAL 1 20

FT CHAIN 21 377 POTENTIAL.

FT DOMAIN 116 256 EJACULATORY BULB SPECIFIC PROTEIN I.

FT DOMAIN 159 377 PRO-RICH.

FT VARIANT 84 84 GLY-RICH.

SQ SEQUENCE 377 AA; 37819 MM; D020F027BD96B221 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 377;

Best Local Similarity 83.3%; Pred. No. 62;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPM 6

DB 261 GAVVPM 266

```

RESULT 6
AROA_BACSU STANDARD: PRT: 428 AA.
ID AROA_BACSU
AC P20691:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSPS).
GN AROE
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87191994; PubMed=3106153;
RA Henner D.J., Band L., Flagg G., Chen E.;
RT "The organization and nucleotide sequence of the Bacillus subtilis
  hsh, tyrA and aroE genes.";
RL Gene 49:147-152(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
MEDLINE=38044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
  Azevedo V., Berrtero M.G., Bessieres P., Bolotin A., Borcher S.,
  Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
  Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
  Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
  Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
  Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
  Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
  Giuseppe G., Guy B.J., Haga K., Halech E.J., Grandi G.,
  Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
  Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
  Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
  Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
  Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
  Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
  Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
  Paro V., Pohl T.M., Portelle D., Portwollik S., Prescott A.M.,
  Prescan E., Pujic P., Purnelle B., Rapoport G., Rey W., Reynolds S.,
  Rieger M., Rivolta C., Roeha E., Roche B., Rose M., Sadleir Y.,
  Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
  Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
  Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
  Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
  Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
  Viari A., Wambut R., Wedler E., Wedler H., Weltenegeger T.,
  Winters P., Wipit A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
  Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
  subtilis.";
RL Nature 390:249-256(1997).
-i CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
  phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -i PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
  sixth step.
CC -i SUBUNIT: Monomer (By similarity).
CC -i SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -i SIMILARITY: Belongs to the EPSP synthase family.
CC -i SIMILARITY: Belongs to the EPSP synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL: M80245; AAA20869.1;

```

```

DR EMBL: Z99115; CAB14176.1;
DR PIR: C26532; C26532.
DR Subtilisin; BG10294; aroE.
DR HAMAP: MF_00210; -; 1.
DR InterPro: IPR006264; AroA.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR TIGRFAMs: TIGR01356; aroA; 1.
DR PROSITE: PS00104; EPSP-SYNTASE_1; 1.
DR PROSITE: PS00885; EPSP-SYNTASE_2; 1.
DR Aromatic amino acid biosynthesis; transferase; Complete proteome.
SQ SEQUENCE 428 AA; 45240 MW; DE3F7B96E761CB40 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 428;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVPN 6
Db 241 GAVPN 246

RESULT 7
ARAB_SALTI STANDARD: PRT: 568 AA.
ID ARAB_SALTI
AC P58542;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE L-ribulokinase (EC 2.7.1.16).
GN ARAB OR STY0120 OR T0107.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
  Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
  Felwell T., Hamlin N., Haque A., Hen T.T., Holroyd S., Jagers K.,
  Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
  Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
  Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=TY2 / ATCC 700931;
MEDLINE=22511367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G., III, Mayhew G.F., Rose D.J.,
  Burdall V., Kodoyianni V., Schwartz D.C., Blatter F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
  and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
-i CATALYTIC ACTIVITY: ATP + L-ribulose = ADP + L-ribulose 5-
  phosphate.
CC -i PATHWAY: L-arabinose catabolism; second step.
CC -i SIMILARITY: Belongs to the ribulokinase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC

```

DR EMBL: AL627265; CAD01260.1; -
 DR HAMAB: AE016834; AA067839.1; -
 DR HAMAB: MF_00520; -; 1.
 DR InterPro: IPR000577; FGGY_kin.
 DR InterPro: IPR005929; L-ribulokin.
 DR Pfam: PF00370; FGGY_1.
 DR Pfam: PF02782; FGGY_C1.
 DR TIGRPFAM: TIGR01234; L-ribulokinase; 1.
 DR Transferase: Kinase; Arabinose catabolism; Complete proteome.
 KW INIT_MER 0
 FT SEQUENCE 568 AA; 61643 MW; A1515DB7249A0F75 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 1; Length 568;
 Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPN 6
 Db 321 GSVVPPN 326

RESULT 8
 ARAB_SALTY STANDARD; PRT; 568 AA.
 AC P06188;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE L-ribulokinase (EC 2.7.1.16).
 GN ARAB OR STM0103.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=602.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85232044; PubMed=2989100;
 RA Lin H.-C., Lei S.-P., Wilcox G.;
 RT "The arabid operon of Salmonella typhimurium ltr2. I. Nucleotide
 RT sequence of arab and primary structure of its product,
 RT ribulokinase.";
 RL Gene 34:111-122(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ribulose = ADP + L-ribulose 5-
 CC phosphate.
 CC -1- PATHWAY: L-arabinose catabolism; second step.
 CC -1- SIMILARITY: Belongs to the ribulokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M11047; AAA27023.1; -
 DR EMBL: AE008698; AAL19067.1; -
 DR PIR: A24984; A24984.
 DR StyGene: SG10013; arab.
 DR HAMAB: MF_00520; -; 1.
 DR InterPro: IPR000577; FGGY_kin.

DR InterPro: IPR005929; L-ribulokin.
 DR Pfam: PF00370; FGGY_1.
 DR Pfam: PF02782; FGGY_C1.
 DR TIGRPFAM: TIGR01234; L-ribulokinase; 1.
 KW Transferase: Kinase; Arabinose catabolism; Complete proteome.
 FT INIT_MER 0
 SO SEQUENCE 568 AA; 61620 MW; B1D838C37BCA134 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 1; Length 568;
 Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPN 6
 Db 321 GSVVPPN 326

RESULT 9
 VDRH_BP186 STANDARD; PRT; 66 AA.
 AC P21680;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE DHR protein.
 GN DHR OR CP78.
 OS Bacteriophage 186.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC P2-like viruses.
 OX NCBI_Taxid=29252;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89199647; PubMed=2704042;
 RA Richardson H., Puspurs A., Egan J.B.;
 RT "Control of gene expression in the p2-related temperate coliphage
 RT 186. VI. Sequence analysis of the early lytic region.";
 RL J. Mol. Biol. 206:251-255(1989).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=89199651; PubMed=2704043;
 RA Richardson H., Egan J.B.;
 RT "DNA replication studies with coliphage 186. II. Depression of host
 RT replication by a 186 gene.";
 RL J. Mol. Biol. 206:59-68(1989).
 CC -1- FUNCTION: INVOLVED IN THE DEPRESSION OF HOST DNA REPLICATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X15001; CAA33108.1; -
 DR EMBL: U32222; AAC34180.1; -
 DR PIR: S03595; S03595.
 DR Early protein.
 KW SEQUENCE 66 AA; 7531 MW; 64E93005DA7C0F7 CRC64;

Query Match
 Best Local Similarity 87.1%; Score 27; DB 1; Length 66;
 Matches 4; Conservativity 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPN 6
 Db 11 GAMIPN 16

RESULT 10
 H2A2_SCHPO STANDARD; PRT; 130 AA.
 ID H2A2_SCHPO
 AC P04910;

DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A-beta (H2A.2).
 GN HTA2 OR SPAC19G12.06c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66310796; PubMed=3018512;
 RA Choe J., Schuster T., Grunstein M.;
 RT "Organization, primary structure, and evolution of histone H2A and
 RT H2B genes of the fission yeast Schizosaccharomyces pombe.";
 RL Mol. Cell. Biol. 5:3261-3269(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66135992; PubMed=4092687;
 RA Matsunoto S., Yanagida M.;
 RT "Histone gene organization of fission yeast: a common upstream
 RT sequence.";
 RL EMBO J. 4:3531-3538(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Wells C., Fuchs W., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules
 CC each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
 CC bp of DNA.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- SIMILARITY: Belongs to the histone H2A family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M1500; AAA35310.1; -
 CC EMBL: X05221; CAA28849.1; -
 CC EMBL: Z97209; CAB10117.1; -
 CC DR PIR: C27399; HS2PA3.
 CC GeneDB_Spombe; SPAC19G12.06c; -

DR InterPro: IPR004822; Histone_core.
 DR InterPro: IPR002119; Histone_H2A.
 DR Pfam: PF00125; histone.1.
 DR PRINTS: PR00620; HISTONEH2A.
 DR ProDom: PD000522; HISTONEH2A.
 DR SMART: SM00414; H2A.1.
 DR PROSITE: PS00046; HISTONE_H2A.1.
 DR Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
 DR Multigene family.
 FT INIT MET
 FT
 SQ SEQUENCE 130 AA; 13645 MW; 1BD5C86BC8692A04 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 130;
 Best local similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 GY 1 GAVPN 6
 DB 106 GGAVPN 111
 RESULT 11
 ID H2A1_SCHPO STANDARD; PRT; 131 AA.
 AC P04909;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A-alpha (H2A.1).
 GN HTA1 OR SFCC622.08c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66310796; PubMed=3018512;
 RA Choe J., Schuster T., Grunstein M.;
 RT "Organization, primary structure, and evolution of histone H2A and
 RT H2B genes of the fission yeast Schizosaccharomyces pombe.";
 RL Mol. Cell. Biol. 5:3261-3269(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66135992; PubMed=4092687;
 RA Matsunoto S., Yanagida M.;
 RT "Histone gene organization of fission yeast: a common upstream
 RT sequence.";
 RL EMBO J. 4:3531-3538(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs W., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe".
RL Nature 413:871-880(2002).
CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules
CC each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC bp of DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the histone H2A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M1494; AAA35311.1; -;
CC EMBL: X05230; CAA28848.1; -;
CC EMBL: AL033127; CAA21864.1; -;
CC PIR: B27399; HSZPA2.
CC GeneDB_Spomb; SPC622.08c; -;
CC DR InterPro: IPR004822; Histone_H2A.
CC DR InterPro: IPR002119; Histone_H2A.
CC DR Pfam: PF00125; histone.1.
CC DR PRINTS: PR00620; HISTONEH2A.
CC DR ProDom: PD000522; Histone_H2A; 1.
CC DR SMART: SM00414; H2A; 1.
CC DR PROSITE: PS00046; HISTONE_H2A; 1.
CC KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
CC MW Multigene family.
CC FT INIT_MET 0
CC FT LENGTH 123
CC FT CONFIDENT 123
CC SQ SEQUENCE 131 AA; 13747 MW; DIAF80F580C3273 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 131;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAVVPN 6
1 |||||
DB 106 GGVPVN 111

RESULT 12
MLIA_XENLA
ID MLIA_XENLA STANDARD; PRT; 153 AA.
AC P51048;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 01-OCT-1996 (Rel. 34; Last annotation update)
DE Melatonin receptor type 1A X2.0 (MEL-1A-R) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodidae; Xenopus.
CC NCBI_TaxID=8355;
CC RX MEDLINE=96073557; PubMed=7576645;
CC RA Reppert S.M., Weaver D.R., Cassone V.M., Godson C.,
CC RA Koltowski L.F. Jr.,
CC "Melatonin receptors are for the birds: molecular analysis of two
CC receptor subtypes differentially expressed in chick brain.";
CC Neuron 15:1003-1015(1995).
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR MELATONIN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY PERTUSSIS TOXIN SENSITIVE G PROTEINS
CC THAT INHIBITS ADENYLYLATE CYCLASE ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U01826; AAA92500.1; -;
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR00237; GPCRHHODOPSN.
CC DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; PARTIAL.
CC DR PROSITE: PS00262; G-PROTEIN_RECEP_FL2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT NON_TER 1
CC FT DOMAIN 1
CC FT TRANSMEM 12
CC FT DOMAIN 13
CC FT TRANSMEM 33
CC FT DOMAIN 34
CC FT TRANSMEM 57
CC FT TRANSMEM 58
CC FT TRANSMEM 78
CC FT TRANSMEM 79
CC FT TRANSMEM 113
CC FT TRANSMEM 133
CC FT TRANSMEM 134
CC FT TRANSMEM 145
CC FT TRANSMEM 146
CC FT NON_TER 153
CC FT NON_TER 153
CC SQ SEQUENCE 153 AA; 17665 MW; ED0421422BE20F27 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 153;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPN 6
1 |||||
DB 27 GATLPN 32

RESULT 13
KAD_RICPR
ID KAD_RICPR STANDARD; PRT; 213 AA.
AC Q9ZCS6;
DT 30-MAY-2000 (Rel. 39; Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase).
GN ADK OR RP638.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
CC NCBI_TaxID=782;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Madrid E;
CC RX MEDLINE=99039499; PubMed=9823893;
CC RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
CC Sichterich-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
CC Eriksson A.-S., Winkler H.H., Kurland C.G.,
CC "The genome sequence of Rickettsia prowazekii and the origin of
CC mitochondria".
CC RT Nature 396:133-140(1998).
CC RL Nature 396:133-140(1998).
CC -1- FUNCTION: THIS SMALL, UBIQUITOUS ENZYME IS ESSENTIAL FOR
CC MAINTENANCE AND CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the adenylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: AJ235272; CAA15078.1;
DR PIR: D71669; D71669.
DR HSSP: P07170; IAKY.
DR HAMAP: ME_00235; -. 1.
DR InterPro: IPR006259; Adenyl_kin_sub.
DR InterPro: IPR000850; Adenylate_kin.
DR Pfam: PF00406; ADK; 1.
DR Pfam: PF05191; ADK_1ld; 1.
DR PRINTS: PR00094; ADENYLTKINASE.
DR PRODOM: PD000657; Adenylate_kin; 1.
DR TIGRFAMs: TIGR01351; adk; 1.
DR PROSITE: PS00113; ADENYLATE_KINASE; 1.
DR TRANSFERASE: Kinase; ATP-binding; Complete proteome.
DR NP_BIND: 7 15 ATP (BY SIMILARITY).
SO SEQUENCE 213 AA; 24503 MW; 23CBD06A1BECC19 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 213;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAYVNP 6
Db 56 GALPN 61

RESULT 14
PRPD_ECO57
ID PRPD_ECO57 STANDARD; PRT; 482 AA.
AC 08X693;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-methylcitrate dehydratase (EC 4.2.1.79).
GN PRPD OR 20429 OR ECS0387.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Pena N.T., Plunkett G., Ili, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boulton A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Ikida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RT DNA Res. 8:11-22(2001).
RN [3]
RP FUNCTION: Catalyzes the dehydration of 2-methylcitrate to 2-
RP methyl-cis-aconitate (By similarity).
CC CATALYTIC ACTIVITY: 2-hydroxybutane-1,2,3-tricarboxylate = (2)-
CC but-2-ene-1,2,3-tricarboxylate + H(2)O.
CC COFACTOR: Contains one 2Fe-2S cluster (By similarity).
CC PATHWAY: Propionate catabolism; 2-methylcitric acid cycle; third
CC step.
CC SUBUNIT: Monomer (By similarity).
CC SIMILARITY: BELONGS TO THE MMGE / PRPD FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb.sib.ch/announce/
CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: AE005212; AAG54682.1;
DR EMBL: AP002551; BAB33810.1; -.
DR PIR: G90677; C90677.
DR PIR: F83527; F83527.
DR InterPro: IPR005656; MMGE_PRPD.
DR Pfam: PF03972; MMGE_PRPD; 1.
DR Lyase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S; Complete proteome.
DR INIT MET 0 BY SIMILARITY.
SO SEQUENCE 482 AA; 53820 MW; 19714CD931C08227 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 482;
Best Local Similarity 83.3%; Pred. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAYVNP 6
Db 66 GTVAVPN 71

RESULT 15
PRPD_ECOLI
ID PRPD_ECOLI STANDARD; PRT; 482 AA.
AC P77243;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-methylcitrate dehydratase (EC 4.2.1.79).
GN PRPD OR B0334.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blattner F.R., Plunkett G., Ili, Bloch C.A., Pena N.T., Burland V.,
RA Riley M.F., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1233-1238(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Nemeth A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RX Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND SEQUENCE OF 1-12.
RX MEDLINE=21642584; PubMed=11782506;
RA Blank L., Green J., Guest J.R.;
RT "Acnc of Escherichia coli is a 2-methylcitrate dehydratase (Prpd) that
RT can use citrate and isocitrate as substrates."
RT Microbiology 148:133-146(2002).
RN [4]
RP FUNCTION: Catalyzes the dehydration of 2-methylcitrate to 2-
RP methyl-cis-aconitate. Also seems to be responsible for the
CC residual aconitase activity of the acnAB-null strain.
CC CATALYTIC ACTIVITY: 2-hydroxybutane-1,2,3-tricarboxylate = (2)-
CC but-2-ene-1,2,3-tricarboxylate + H(2)O.
CC COFACTOR: Contains one 2Fe-2S cluster.
CC PATHWAY: Propionate catabolism; 2-methylcitric acid cycle; third
CC step.
CC SUBUNIT: Monomer.
CC SIMILARITY: BELONGS TO THE MMGE / PRPD FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE000140; AAC73437.1; -.
DR EMBL: U73857; AAB18038.1; -.
DR PIR: F64760; F64760.
DR EcoGene: EG13603; PRPD.
DR InterPro: IPR005656; MmgE_PripD.
DR Pfam: PF03972; MmgE_PripD.1.
KW lyase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S; Complete proteome.
FT INIT_MET 0
SQ SEQUENCE 482 AA; 53820 MW; A7AA5CD5391B82C5 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 482;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAVVPN 6
I I I I I
Db 66 GTVVPN 71

Search completed: September 4, 2003, 21:01:43
Job time : 7.20588 secs


```

DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DB Putative thioredoxin.
OS OSJNB0091p11.16.
OC Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Beil C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tstirn T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
RA VanAken S.E., Uterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.,
RT "Oryza sativa chromosome 3 BAC OSJNB0091p11 genomic sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC073556; AAL84306.1; -
DR Gramene; 085708; -
DR InterPro: IPR006663; Thioredox_dom2.
DR InterPro: IPR006663; Thioredox_dom2.
DR Pfam: PF00085; ThioRed_1.
SQ SEQUENCE 234 AA; 26253 MW; B6FC2C8293EB11C CRC64;

Query Match 100.0%; Score 31; DB 10; Length 234;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
DB 227 GAVPN 232

RESULT 3
O45534 PRELIMINARY; PRT; 468 AA.
AC O45534;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE F49A5.7 protein.
GN F49A5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81542; CAB04419.1; -.
DR HSSP; P05451; 1DD.
DR WormPep; F49A5.7; CE16070.
DR InterPro: IPR001304; LECTIN_C.
DR Pfam; SM00034; LECTIN_C; 2.
DR SMART; PF00059; LECTIN_C; 3.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_2; 2.
SQ SEQUENCE 468 AA; 52464 MW; 71E731E29EF07118 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 468;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GAVPN 6
DB 401 GAVPN 406

RESULT 4
O04236 PRELIMINARY; PRT; 72 AA.
ID O04236;
AC O04236;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Potassium channel (Fragment).
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=97272307; PubMed=9114073;
RA Hoth S., Dreyer I., Dietrich P., Becker D., Mueller-Roeber B.,
RA Hedrich R.;
RT "Molecular basis of plant-specific acid activation of K+ uptake
RT channels.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4806-4810(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Philipp K.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09749; CAA70896.1; -.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR Pfam; PF00520; Ion_trans; 1.
KW Ionic channel; Transmembrane.
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8259 MW; D8B6A36D43C2F00F CRC64;

Query Match 96.8%; Score 30; DB 10; Length 72;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
DB 5 GAVPN 10

RESULT 5
O04241 PRELIMINARY; PRT; 72 AA.
ID O04241;
AC O04241;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Potassium channel (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS51311B; TISSUE=Coleoptile;
RX MEDLINE=97272307; PubMed=9114073;
RA Hoth S., Dreyer I., Dietrich P., Becker D., Mueller-Roeber B.,
RA Hedrich R.;
RT "Molecular basis of plant-specific acid activation of K+ uptake
RT channels.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4806-4810(1997).
RN [2]

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN-L51311B; TISSUE=Coleoptile;
RL  Philippiar K.;
DR  Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL: Y09747; CAAT0894.1; -.
DR  InterPro: IPR005821; Ion_trans.
DR  InterPro: IPR001622; K-channel_pore.
DR  Pfam: PF00520; Ion_trans; 1.
KW  Ionic channel; Transmembrane.
FT  NON_TER 1
FT  NON_TER 72
SQ  SEQUENCE 72 AA; 8283 MW; 56B84B6AFE3D5DED CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 10; Length 72;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 GAVVFN 6
    111:11
DB  5 GAVVFN 10

RESULT 6
ID  042199      PRELIMINARY;      PRT;      108 AA.
AC  042199;
DT  01-JAN-1998 (TREMblrel. 05, Created)
DT  01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT  01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE  Beta-2 microglobulin precursor (Fragment).
GN  B2M.
OS  Ictalurus pricei.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC  Ictaluridae; Ictalurus.
OX  NCBI_TaxID=64534;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98435523; PubMed=9745011;
RA  Criscitello M.F., Benedetto R., Antao A., Wilson M.R., Chinchar V.G.,
RA  Miller N.W., Clem L.W., McConnell T.J.
RT  "beta2-microglobulin of ictalurus catfishes.";
RL  Immunogenetics 48:339-343(1998).
DR  EMBL: AF016044; AAC64993.1; -.
DR  HSSP; P01888; IBMG.
DR  InterPro: IPR007110; Ig-like.
DR  InterPro: IPR003597; Ig-cl.
DR  InterPro: IPR003006; Ig_MHC.
DR  Pfam: PF00047; Ig; 1.
DR  SMART; SM00407; Igcl; 1.
DR  PROSITE; PS50835; IG_LIKE; 1.
DR  PROSITE; PS00290; IG_MHC; 1.
KW  Signal.
FT  NON_TER 1
FT  SIGNAL <1 11
FT  CHAIN 12 >108 POTENTIAL.
FT  NON_TER 108 BETA-2 MICROGLOBULIN.
SQ  SEQUENCE 108 AA; 12163 MW; 507EEF13B0A2A9AB CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 13; Length 108;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 GAVVFN 6
    111:11
DB  54 GAVVFN 59

RESULT 7
ID  0931A6      PRELIMINARY;      PRT;      340 AA.
AC  0931A6;
DT  01-DEC-2001 (TREMblrel. 19, Created)

```

```

DT  01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT  01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE  Putative LacI-family transcriptional regulator.
GN  RA0038 OR SMA0078
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Plasmid pSymba (megaplasmid 1).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium.
OX  NCBI_TaxID=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RX  MEDLINE=21396509; PubMed=11481432;
RA  Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA  Barloy-Hubler F., Bowser L., Capela D., Galibert F., Guzy J.,
RA  Gujral M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA  Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA  Yen K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT  "Nucleotide sequence and predicted functions of the entire
RT  Sinorhizobium meliloti pSymba megaplasmid.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR  EMBL: AE007198; AAK64696.1; -.
DR  InterPro: IPR000843; HTH_Lact.
DR  InterPro: IPR001761; Periplase/Lact.
DR  Pfam: PF00356; lact; 1.
DR  Pfam: PF00532; Peripla_BP_like; 1.
DR  SMART; SM00354; HTH_Lact; 1.
DR  PROSITE; PS00356; HTH_LACT_FAMILY; 1.
KW  Plasmid; Hypothetical protein; Complete proteome.
SQ  SEQUENCE 340 AA; 36655 MW; C5722E9EC7FA3208 CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 16; Length 340;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 GAVVFN 6
    111:11
DB  72 GAVVFN 77

RESULT 8
ID  08XTH4      PRELIMINARY;      PRT;      411 AA.
AC  08XTH4;
DT  01-MAR-2002 (TREMblrel. 20, Created)
DT  01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT  01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE  Putative NADH oxidase-related oxidoreductase protein
DE  (EC 1.-.-.-).
GN  RSP0136 OR RS02984.
OS  Ralstonia solanacearum (Pseudomonas solanacearum).
OC  Plasmid megaplasmid.
OC  Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC  Ralstoniaceae; Ralstonia.
OX  NCBI_TaxID=305;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=GMT1000;
RX  MEDLINE=21681879; PubMed=11823852;
RA  Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
RA  Ariat M., Billault A., Brotier P., Camus J.C., Cattolico L.,
RA  Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA  Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA  Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA  Weissenbach J., Boucher C.A.;
RT  "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL  Nature 415:497-502(2002).
DR  EMBL: AL646076; CAD17287.1; -.
DR  InterPro: IPR001155; Oxidored_FMN.
DR  Pfam: PF00724; Oxidored_FMN; 1.
KW  Oxidoreductase; Plasmid; Complete proteome.
SQ  SEQUENCE 411 AA; 44798 MW; 88AF324B7A305FCB CRC64;

```

Query Match 96.8%; Score 30; DB 16; Length 411;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 DB 12 GAVIPN 17

RESULT 9

039279 PRELIMINARY; PRT; 587 AA.
 ID 039279
 AC 039279
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Counterpart of HSV-1 gene UL25 and VZV gene 34.
 GN 36.
 OS Equine herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=10331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NS80567;
 RX MEDLINE=98264497; PubMed=9603335;
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus 4.";
 RN J. Gen. Virol. 79:1197-1203(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NS80567;
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF030027; AAC59553.1;
 DR InterPro; IPR002493; UL25.
 DR Pfam; PF01499; UL25; 1.
 SQ SEQUENCE 587 AA; 63615 MW; B8F2F791P44CC958 CRC64;

Query Match 96.8%; Score 30; DB 12; Length 587;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 DB 413 GAVIPN 418

RESULT 10

08B9C8 PRELIMINARY; PRT; 803 AA.
 ID 08B9C8
 AC 08B9C8;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Rachiplusia ou multiple nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=80366;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bonning B.C., Harrison R.L.;
 RT "The Rachiplusia ou multiple nucleopolyhedrovirus genome sequence."
 RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY145471; AAN28018.1;
 KM Hypothetical protein.
 SQ SEQUENCE 803 AA; 94645 MW; 30D414D7B37EF22E CRC64;

Query Match 96.8%; Score 30; DB 12; Length 803;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 DB 412 GAVIPN 417

RESULT 11

024538 PRELIMINARY; PRT; 807 AA.
 ID 024538
 AC 024538;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Potassium channel (Fragment).
 OS Vicia faba (Broad bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
 OX NCBI_TaxID=3906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cotyledon;
 RA Ache P., Mohlfarth T., Hedrich R., Becker D.;
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Y10579; CA71598.1;
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR00595; CNMP_binding.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001622; K_channel_pore.
 DR InterPro; IPR005820; M_channel_nlg.
 DR Pfam; PF00023; ank; 5.
 DR Pfam; PF00027; CNMP_binding; 1.
 DR Pfam; PF00520; Ion_trans; 1.
 DR PRINTS; PR01415; ANKTRIN.
 DR SMART; SM00248; ANK; 3.
 DR SMART; SM00100; CNMP; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 1.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50042; CNMP_BINDING; 3; 1.
 KW ANK repeat; Ionic channel; Repeat; Transmembrane.
 FT NON_TER 807
 SQ SEQUENCE 807 AA; 92464 MW; D0B01A2CD25A1038 CRC64;

Query Match 96.8%; Score 30; DB 10; Length 807;
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 DB 230 GAVIPN 235

RESULT 12

09FY04 PRELIMINARY; PRT; 830 AA.
 ID 09FY04
 AC 09FY04;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Potassium channel 2.
 GN PTK2.
 OS Populus tremula x Populus tremuloides.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=47664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Phloem; Phloem, and Xylem;
 RA Ache P., Langer K., Regan S., Geiger D., Hedrich R.;
 RT "Potassium dependent wood formation."
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ271447; CAC05489.1;
 DR InterPro; IPR002110; ANK.

```

DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR Pfam: PF00023; ank; 5.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 3.
DR SMART: SM00100; CNMP; 1.
DR PROSITE: PSS0088; ANK_REPEAT; 2.
DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
DR ANK repeat; Ionic channel; Repeat; Transmembrane.
SO SEQUENCE 830 AA; 94186 MW; 26D0458F845BC821 CRC64;

Query Match
Best Local Similarity 83.3%; Score 30; DB 10; Length 830;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
Db 237 GAVPN 242

RESULT 13
O9ZPL4 PRELIMINARY; PRT; 832 AA.
AC O9ZPL4;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Pulvinus inward-rectifying channel for potassium SPIK1.
OS Samanea saman.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Fabales; Fabaceae; Mimosoidae; Ingeae; Samanea.
OX NCBI_TaxID=76910;
RN [1]
RP SEQUENCE FROM N.A.
RA Mosheleion M., Moran N., Hedrich R., Becker D.;
RT "Molecular and biophysical analysis of Samanea saman motor cells.";
RL EMBL: AF099095; AAD16278.1; -
DR HSSP: Q54397; 1BL8.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR Pfam: PF00023; ank; 5.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 2.
DR SMART: SM00100; CNMP; 1.
DR PROSITE: PSS0088; ANK_REPEAT; 3.
DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
DR ANK repeat; Ionic channel; Repeat; Transmembrane.
SO SEQUENCE 832 AA; 95189 MW; EE90E2A51CE4707 CRC64;

Query Match
Best Local Similarity 83.3%; Score 30; DB 10; Length 832;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
Db 255 GAVPN 260

RESULT 14
O9SM12

```

```

ID O9SM12 PRELIMINARY; PRT; 849 AA.
AC O9SM12;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Potassium channel protein ZmK2.
GN ZmK2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA SRAIN-155131B. and cv. Apache; TISSUE=Coleoptile;
RC MEDLINE-20604476; PubMed-10518597;
RX Phillippar K., Fuchs I., Luthen H., Hoth S., Bauer C.S., Haga K.,
RA Thiel G., Jung K., Sandberg G., Bottger M., Becker D., Hedrich R.;
RT "Auxin-induced K+ channel expression represents an essential step in
RT coleoptile growth and gravitropism.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:12186-12191(1999).
DR EMBL: AJ132686; CAB54856.1; -
DR HSSP: P42773; 1IIB.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR Pfam: PF00023; ank; 4.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR SMART: SM00248; ANK; 2.
DR SMART: SM00100; CNMP; 1.
DR PROSITE: PSS0088; ANK_REPEAT; 3.
DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
DR ANK repeat; Ionic channel; Repeat; Transmembrane.
SO SEQUENCE 849 AA; 93951 MW; D62EBA1276A2B07B CRC64;

Query Match
Best Local Similarity 83.3%; Score 30; DB 10; Length 849;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
Db 244 GAVPN 249

RESULT 15
O97017 PRELIMINARY; PRT; 1891 AA.
AC O97017;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE L-type calcium channel alpha-1 subunit.
GN STPCACHL.
OS Stylophora pistillata.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Asterozoenina; Pocilloporidae; Stylophora.
OX NCBI_TaxID=50429;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99148007; PubMed-10023047;
RX Zoccola D., Tambutte E., Senegas-Balas F., Michiels J.F., Failla J.P.,
RA Jaudert J., Allemand D.;
RT "Cloning of a calcium channel alpha1 subunit from the reef-building
RT coral, Stylophora pistillata.";
RL Gene 227:157-167(1999).
DR EMBL: U64465; AAD11470.1; -
DR InterPro: IPR001682; Ca/Na_pore.
DR InterPro: IPR002111; Ca_channel_TrypL.
DR InterPro: IPR002077; Ca_channel.

```

DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR005446; LVDCALpha1.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00520; Ion_trans; 4.
 DR PRINTS; PRO0167; CACHANNEL.
 DR PRINTS; PRO1630; LVDCALPHA1.
 KW Ionic channel; Transmembrane.
 SQ SEQUENCE 1891 AA; 213240 MW; BAB705AEBF9BA134 CRC64;

Query Match 96.8%; Score 30; DB 5; Length 1891;
 Best Local Similarity 83.3%; Pred. No. 7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 ||:||||
 Db 744 GAVVFN 749

Search completed: September 4, 2003, 21:06:39
 Job time : 26.7086 secs

PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 25; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVVPN 5
DB 1 AVVPN 5
RESULT 2
AAE16722
ID AAE16723 standard; peptide: 5 AA.
XX
AC AAE16723;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #3.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypertensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 5 /note="C-terminal amide"
FT
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;

XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 25; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVVPN 5
DB 1 AVVPN 5
RESULT 3
AAE16722
ID AAE16722 standard; peptide: 6 AA.
XX
AC AAE16722;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #2.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypertensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin

XX Claim 24; Page 21; 53pp; English.

PS
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

SQ Sequence 6 AA:

Query Match 100.0%; Score 25; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
|||||
DB 2 AVVPN 6

RESULT 4
AAE16724
ID AAE16724 standard; peptide; 6 AA.
XX
AC AAE16724;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #4.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note- "C-terminal amide"
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGINGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin

PS Claim 24; Page 21; 53pp; English.

XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.

SQ Sequence 6 AA:

Query Match 100.0%; Score 25; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
|||||
DB 2 AVVPN 6

RESULT 5
AAE16726
ID AAE16726 standard; peptide; 7 AA.
XX
AC AAE16726;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #6.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGINGING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for

CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hyperextension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 25; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
|||||
Db 2 AVVPN 6
RESULT 6
AAE16729 standard; peptide; 7 AA.
XX
AC AAE16729;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #9.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
FH Key
FH Modified-site 1 Location/Qualifiers
FT /note= "This residue forms a cyclic structure with Cys
FT at position 7 with copper as a chelating agent"
FT Modified-site 7
FT /note= "This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent"
FT
FT
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PE 30-MAY-2001; 2001WO-US17384.
XX
PF 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mitts TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.

XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hyperextension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived cyclic peptide.
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 25; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
|||||
Db 2 AVVPN 6
RESULT 7
AAE16727 standard; peptide; 8 AA.
XX
AC AAE16727;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
FH Key
FH Modified-site 1 Location/Qualifiers
FT /note= "This residue forms a cyclic structure with Cys
FT at position 7 with copper as a chelating agent"
FT Modified-site 7
FT /note= "This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent"
FT
FT
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PE 30-MAY-2001; 2001WO-US17384.
XX
PF 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
PI Mitts TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin -
XX
PS Claim 24; Page 21; 53pp; English.
XX
PT The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of

CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HBP), VVPN derived peptide.

XX Sequence 8 AA:

Query Match 100.0%; Score 25; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 |||||
 Db 3 AVVPN 7

RESULT 8
 AAE16730
 ID AAE16730 standard; peptide; 8 AA.

AC AAE16730;
 DT 09-APR-2002 (first entry)

XX Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.

XX Hydrolysed elastin peptide; HBP: therapy; cosmetic appearance; cardiant;
 KM vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KM hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KM restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KM cyclic.

OS Ligamentum nuchae.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1

FT /note="this residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"

FT Modified-site 8
 FT /note="this residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"

XX WO200191700-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17384.

XX 30-MAY-2000; 2000US-0580110.

XX 30-MAY-2000; 2000US-0580156.

XX 30-MAY-2000; 2000US-0580893.

XX 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.

XX Mlts TF, Sandberg LB, Jimenez F;

XX WPI; 2002-106259/14.

XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX Claim 24; Page 21; 53pp; English.

CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HBP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HBP), VVPN derived cyclic peptide.

XX Sequence 8 AA:

Query Match 100.0%; Score 25; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 |||||
 Db 3 AVVPN 7

RESULT 9
 ABG52100
 ID ABG52100 standard; Peptide; 27 AA.

AC ABG52100;

DT 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 30748.

XX Human liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KM hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver
 XX Claim 27; SEQ ID No 30748; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolemia which
CC is associated with coronary heart disease. AB647348-AB659930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
18 AVVPN 22
DB 18 AVVPN 22
RESULT 10
ABB32028
ID ABB32028 standard; Peptide: 27 AA.
XX
AC ABB32028;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #4679 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180313.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX
PS Claim 27; SEQ ID NO 14996; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BR 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias

CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVVPN 5
18 AVVPN 22
DB 18 AVVPN 22
RESULT 11
ABB37278
ID ABB37278 standard; Peptide: 27 AA.
XX
AC ABB37278;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #4784 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180313.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver
XX
PS Claim 27; SEQ ID NO 29913; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
11111
Db 18 AVVPN 22

RESULT 12

ABB22567
ID ABB22567 standard; Protein; 27 AA.

AC ABB22567;

DT 23-JAN-2002 (first entry)

DE Protein #4566 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI: 2001-488899/53.

CC Single exon nucleic acid probes for analyzing gene expression in human

CC hearts -

CC Claim 15; SEQ ID NO 24337; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 27 AA;

Query Match 100.0%; Score 25; DB 22; Length 27;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
11111
Db 18 AVVPN 22

RESULT 13

AAM57980

ID AAM57980 standard; Protein; 27 AA.

XX AAM57980;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30085.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KM epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI: 2001-483446/52.

CC Single exon nucleic acid probes for analyzing gene expression in human

CC brains -

CC Example 4; SEQ ID NO: 30085; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX SQ Sequence 27 AA;

Query Match 100.0%; Score 25; DB 22; Length 27;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
11111
Db 18 AVVPN 22

RESULT 14

AAM70410

ID AAM70410 standard; Protein; 27 AA.

AC AAM70410;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30716.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

```

PN WO200157276-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI: 2001-488900/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX
PS Example 4; SEQ ID NO: 30716; 658bp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX
SQ Sequence 27 AA:
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAVPN 5
DB 18 AAVPN 22

```

RESULT 15
AAM18247
ID AAM18247 standard; Protein; 27 AA.
AC AAM18247;
XX
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #4681 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

```

XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI: 2001-488901/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX
PS Claim 27; SEQ ID NO 23073; 487bp; English.
XX
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 27 AA:
Query Match 100.0%; Score 25; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAVPN 5
DB 18 AAVPN 22

```

Search completed: September 4, 2003, 21:00:31
Job time : 25.7353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 7.79412 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-58
Perfect score: 25
Sequence: 1 AVVPN 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	149	2	regulatory protein
2	25	100.0	161	2	probable oxidoredu
3	25	100.0	229	2	hypothetical prote
4	25	100.0	242	2	DJ1100H13.4 protei
5	25	100.0	265	1	nonstructural prot
6	25	100.0	284	1	deoxyribonuclease
7	25	100.0	292	2	transcription regu
8	25	100.0	292	2	transcription regu
9	25	100.0	302	2	protein kinase cdc
10	25	100.0	334	2	hypothetical prote
11	25	100.0	334	2	hypothetical prote
12	25	100.0	351	2	conserved hypotet
13	25	100.0	363	1	probable hexosyltr
14	25	100.0	364	2	dihydrodipicolinat
15	25	100.0	364	2	dihydrodipicolinat
16	25	100.0	365	2	probable dihydrodi
17	25	100.0	387	1	MG372 homolog G12
18	25	100.0	408	2	hypothetical prote
19	25	100.0	418	2	nickel-cobalt resi
20	25	100.0	442	2	YopN [Imported] -
21	25	100.0	442	2	type III secretion
22	25	100.0	467	2	hypothetical prote
23	25	100.0	468	2	hypothetical prote
24	25	100.0	481	2	asparaginyl-tRNA s
25	25	100.0	503	2	probable membrane
26	25	100.0	513	2	hypothetical prote
27	25	100.0	589	2	hypothetical prote
28	25	100.0	604	2	probable splicing
29	25	100.0	614	2	phaseolin G-box bl

30	25	100.0	620	2	T50150	yeast nrd1-like pr
31	25	100.0	623	2	T52293	MYC-related DNA bi
32	25	100.0	624	2	JC5471	regulatory protein
33	25	100.0	625	2	S18420	regulatory protein
34	25	100.0	642	2	T10861	phaseolin G-box bl
35	25	100.0	663	2	AE0384	cytochrome O ubiqu
36	25	100.0	693	2	AI0996	4-alpha-glucanotra
37	25	100.0	735	2	AE2169	cation transport A
38	25	100.0	850	2	T23478	hypothetical prote
39	25	100.0	954	2	A87431	regulatory protein
40	25	100.0	964	2	T41547	hypothetical prote
41	25	100.0	1221	2	T35227	probable nitrate r
42	25	100.0	1360	2	T12064	DNA binding protei
43	24	96.0	65	2	H84098	hypothetical prote
44	24	96.0	68	2	G97737	hypothetical prote
45	24	96.0	72	2	T12130	potassium channel

ALIGNMENTS

RESULT 1
S02164
regulatory protein flae - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 04-Mar-2000
C:Accession: S02164
R:Kaplan, J.B.; Dingwall, A.; Bryan, R.; Chamber, R.; Shapiro, L.
J. Mol. Biol. 205, 71-83, 1989
A:Title: Temporal regulation and overlap organization of two Caulobacter flagellar ge
A:Reference number: S02164; MUID:89178645; PMID:2648000
A:Accession: S02164
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-149 <KAP>
C:Genetics:
A:Gene: flae
C:Superfamily: Caulobacter crescentus regulatory protein flae
C:Keywords: transcription regulation

Query Match 100.0%; Score 25; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
Db 47 AVVPN 51

RESULT 2
C95858
probable oxidoreductase protein SMB20131 [Imported] - Sinorhizobium meliloti (strain
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 22-Oct-2001
C:Accession: C95858
R:Finan, T.M.; Weidner, S.; Wong, K.; Bhramester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: C95858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <KUR>
A:Cross-References: GB:AL591985; PIDN:CAC48531.1; PID:g15140003; GSPDB:GNO0167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Calbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation
 C:Genetics:
 A:Gene: Smb20131
 A:Genome: plasmid
 C:Superfamily: isocoumarin 1-oxidoreductase alpha chain; ferredoxin [2Fe-2S] homology
 C:Keywords: 2Fe-2S; metalloprotein
 F:11,66,49,61/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 100.0%; Score 25; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 Db 17 AVVPN 21

RESULT 3
 F75078
 hypothetical protein PAB1691 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000
 C:Accession: F75078
 R:anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: F75078

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <RAM>

A:Cross-references: GB:RJ248286; GB:AL096836; NID:95458366; PIDN:CA849931.1; PID:9545844

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1691

C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1869

Query Match 100.0%; Score 25; DB 2; Length 229;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 Db 31 AVVPN 35

RESULT 4
 H59433

DJ1100H13.4 protein [imported] - human

C:Species: Homo sapiens (man)

C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 23-Sep-2002

C:Accession: H59433

R:Smith, M.

Submitted to Genbank, March 2001

A:Reference number: H59433

A:Accession: H59433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-242 <SMIT>

A:Cross-references: GB:CA017688; PTD:913397832; PIDN:CA017688.2

Query Match 100.0%; Score 25; DB 2; Length 242;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 Db 142 AVVPN 146

RESULT 5
 MNV00V
 nonstructural protein NS - Rift Valley fever virus

C:Species: Rift Valley fever virus

A:Note: host (mosquito); Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999

C:Accession: C38552; S11743

R:Giorgi, C.; Accardi, L.; Nicoletti, L.; Gro, M.C.; Takehara, K.; Hilditch, C.; Mori

Vitology 180; 738-753, 1991

A:Title: Sequences and coding strategies of the S RNAs of Toscana and Rift Valley fev

A:Reference number: A38552; MUID:91111992; PMID:1846496

A:Accession: C38552

A:Molecule type: genomic RNA

A:Residues: 1-265 <GIO>

A:Cross-references: GB:X53771; NID:961928; PIDN:CAA37788.1; PID:961929

C:Genetics:

A:Map position: segment S

C:Superfamily: Punta Toro virus nonstructural protein NS

C:Keywords: nonstructural protein

Query Match 100.0%; Score 25; DB 1; Length 265;
 Best Local Similarity 100.0%; Pred. No. 11e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 Db 208 AVVPN 212

RESULT 6
 JC2526

deoxyribonuclease I (EC 3.1.21.1) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JC2526

R:Peitsch, M.C.; Immler, M.; French, L.E.; Tschopp, J.

Biochem. Biophys. Res. Commun. 207, 62-68, 1995

A:Title: Genomic organisation and expression of mouse deoxyribonuclease I.

A:Reference number: JC2526; MUID:95160718; PMID:7857306

A:Accession: JC2526

A:Molecule type: mRNA

A:Residues: 1-284 <PEI>

A:Cross-references: EMBL:000478; NID:9437052; PIDN:AAA03710.1; PID:9437053

C:Superfamily: deoxyribonuclease I

C:Keywords: actin binding; calcium binding; hydrolase

Query Match 100.0%; Score 25; DB 1; Length 284;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 Db 246 AVVPN 250

RESULT 7
 AB1724

transcription regulators (LysR family) homolog lin2335 [imported] - Listeria innocua

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1724

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fshl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kunu, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schluter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1724

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-292 <GLA>

A:Cross-references: GB:AL592022; PIDN:CA097562.1; PID:916414858; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lln2335

Query Match

Best Local Similarity 100.0%; Score 25; DB 2; Length 292;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 AVPN 5
|||||
DB 239 AVPN 243

RESULT 8

A11353

transcription regulators (LysR family) homolog lmo2233 [imported] - *Listeria monocytogenes*C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: A11353

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Meok, C.; Schlueder, T.; Simoes, N.; Tisseret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

C:Accession: A11353

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-292 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAD00311.1; PID:g16411703; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo2233

Query Match

Best Local Similarity 100.0%; Score 25; DB 2; Length 292;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 AVPN 5
|||||
DB 239 AVPN 243

RESULT 9

T11715

protein kinase cdc2a (EC 2.7.1.-), cyclin-dependent - garden snapperagon

C:Species: *Anthriscum majus* (garden snapperagon)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T11715

R:Podert, P.R.; Gaudin, V.; Lunness, P.; Coen, E.S.; Doonan, J.H.

Plant Cell 8, 1465-1476, 1996

A:Title: Distinct classes of cdc2-related genes are differentially expressed during the

A:Reference number: 218695; MUID:96434535; PMID:8837502

A:Accession: T11715

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-302 <FOB>

A:Cross-references: EMBL:X97637; NID:g1321671; PIDN:CA66233.1; PID:g1321672

C:Genetics:

A:Note: cdc2a

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: phosphotransferase

Query Match

Best Local Similarity 100.0%; Score 25; DB 2; Length 302;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 AVPN 5
|||||
DB 259 AVPN 263

RESULT 10

AG2172

hypothetical protein all2934 [imported] - *Nostoc* sp. (strain PCC 7120)C:Species: *Nostoc* sp. PCC 7120A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AG2172

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriku Nakazaki, N.; Shimpoto, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG2172

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA074633.1; PID:g17132028; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

C:Superfamily: *Escherichia coli* glutaminase homolog ynfHQuery Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 334;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 AVPN 5
|||||
DB 277 AVPN 281

RESULT 11

E86794

hypothetical protein ynfJ [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL140C:Species: *Lactococcus lactis* subsp. *lactis*

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: E86794

R:Botolin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*

A:Reference number: AB6625; MUID:21235186; PMID:11337471

A:Accession: E86794

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <STO>

A:Cross-references: GB:AF005176; PID:g12724340; PIDN:AAK05455.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: ynfJf

Query Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 334;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 AVPN 5
|||||
DB 249 AVPN 253

RESULT 12

C82755

conserved hypothetical protein XF0848 [imported] - *Xylella fastidiosa* (strain 9a5c)C:Species: *Xylella fastidiosa*

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: C82755

R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq

Native 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: AB2515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: C82755

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <SIM>

A:Cross-references: GB:AE003924; GB:AE003849; NID:g9105750; PIDN:AAF83658.1; GSPDB:GN001
 A:Experimental source: strain 945C
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Biondes, M.R.S.; Bueno, M.R.P.; Canarço, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohn
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, M.H.; Vialada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0848

Query Match 100.0%; Score 25; DB 2; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AWPVN 5
 |||||
 DB 159 AWPVN 163

RESULT 13

C72590

probable hexosyltransferase (EC 2.4.1.-) APE1191 [similarity] - Aeropyrum pernix (strain
 N:Alternate names: probable capm protein

C:Species: Aeropyrum pernix
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000

C:Accession: C72590

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10362966

A:Accession: C72590

A:Molecule type: DNA
 A:Residues: 1-363 <KAW>

A:Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAAB0177.1; PID:g5104863
 A:Experimental source: strain K1

C:Genetics:
 A:Gene: APE1191
 C:Superfamily: probable hexosyltransferase YLXN

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 100.0%; Score 25; DB 1; Length 363;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AWPVN 5
 |||||
 DB 166 AWPVN 170

RESULT 14

S46304

dihydrodipicolinate synthase (EC 4.2.1.52) precursor [similarity] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S46304

R:Vauterin, M.; Jacobs, M.
 Plant Mol. Biol. 25, 545-550, 1994

A:Title: Isolation of a poplar and an Arabidopsis thaliana dihydrodipicolinate synthase
 A:Reference number: S46304; MUID:94325478; PMID:8049377

A:Accession: S46304
 A:Molecule type: mRNA

A:Residues: 1-364 <VAD>
 A:Cross-references: EMBL:X72971; NID:g530004; PIDN:CAB45642.1; PID:g5139329

C:Genetics:
 A:Genome: nuclear

C:Function:
 A:Description: catalyzes the formation of dihydrodipicolinate from L-aspartate 4-semi

A:Pathway: lysine biosynthesis

A:Note: first step in diaminopimelate and lysine biosynthesis; sensitive to lysine in

C:Superfamily: dihydrodipicolinate synthase

C:Keywords: carbon-oxygen lyase; chloroplast; hydro-lyase; lysine biosynthesis

F:1-38/Domain: transit peptide (chloroplast) #status predicted <TNP>

F:39-364/Product: dihydrodipicolinate synthase #status predicted <MAT>

F:221/Active site: Lys (covalent pyruvate-binding) #status predicted

Query Match 100.0%; Score 25; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AWPVN 5
 |||||
 DB 39 AWPVN 43

RESULT 15

T47899

dihydrodipicolinate synthase precursor - Arabidopsis thaliana
 N:Alternate names: protein T4C21.290

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000

C:Accession: T47899

R:Choisme, N.; Robert, C.; Brothier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.;
 W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479
 A:Accession: T47899

A>Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-364 <CHO>
 A:Cross-references: EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21
 C:Genetics:

A:Map position: 3
 A:Introns: 29/2; 56/2

A:Note: T4C21.290
 C:Superfamily: dihydrodipicolinate synthase

Query Match 100.0%; Score 25; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AWPVN 5
 |||||
 DB 39 AWPVN 43

Search completed: September 4, 2003, 21:10:48
 Job time : 7.79412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11: Search time 4.33824 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-58
Perfect score: 25
Sequence: 1 AVVPN 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	265	1 VNST_RVFPVZ	P21698 rift valley
2	25	100.0	281	1 DRNL_RABIT	O18998 oryctolagus
3	25	100.0	284	1 DRNL_MOUSE	P49183 mus musculu
4	25	100.0	294	1 CC2A_ANTMA	Q38772 antirrhinum
5	25	100.0	294	1 CDC2_CHERU	P93101 cheropodium
6	25	100.0	364	1 DAP1_ARATH	Q912X6 arabidopsis
7	25	100.0	365	1 DAP2_ARATH	Q912C8 arabidopsis
8	25	100.0	387	1 TH11_MYCPN	P75328 mycoplasma
9	25	100.0	418	1 CNRC_ALCEU	P37974 alcaligenes
10	25	100.0	466	1 SYN_VIRCH	Q9K5F9 vibrio chol
11	25	100.0	466	1 SYN_VIRBP	Q87H55 vibrio para
12	25	100.0	466	1 SYN_VIRBU	O8d4f0 vibrio vuln
13	25	100.0	624	1 NIRA_AZOBI	P54929 azospirillum
14	25	100.0	625	1 NIRA_AZOBR	P30667 azospirillum
15	25	100.0	954	1 FLEY_CANDR	P15345 caulobacter
16	25	100.0	964	1 YOKA_SCHPO	O74522 schizosach
17	25	100.0	1360	1 GLI1_XENLA	O91690 xenopus lae
18	25	100.0	152	1 ML12_BRARE	P51047 brachydanio
19	25	100.0	153	1 ML11_BRARE	P51046 brachydanio
20	25	100.0	154	1 ML1A_BRARE	O02781 sus scrofa
21	25	100.0	169	1 V55_BPT7	O02781 sus scrofa
22	25	100.0	232	1 6PGL_HAEIN	O57039 haemophilus
23	25	100.0	242	1 PEPE_ANASP	P58493 anabaena sp
24	25	100.0	249	1 Y084_METJA	O06032 methanococ
25	25	100.0	257	1 Y084_METJA	O02769 bacillus su
26	25	100.0	283	1 Y084_METJA	P37524 bacillus su
27	25	100.0	318	1 KPR2_YEAST	P38620 saccharomyc
28	25	100.0	319	1 KPR2_YEAST	O97620 clostridium
29	25	100.0	337	1 G3P_METTH	O27090 methanobact
30	25	100.0	338	1 G3P_METTH	P19314 methanobact
31	25	100.0	346	1 G3P_METTH	P19315 methanobact
32	25	100.0	348	1 G3P_METTH	P49288 gallus gall
33	25	100.0	353	1 ML1A_CHICK	P49285 gallus gall

34	24	96.0	366	1 ML1A_SHEEP	P48040 ovis aries
35	24	96.0	379	1 Y528_SYNY3	O5518 synchocyst
36	24	96.0	380	1 G1GC_BACSU	P39122 bacillus su
37	24	96.0	420	1 ML1C_XENLA	P49219 xenopus lae
38	24	96.0	420	1 SECT_GUTTH	P28527 guilliardea
39	24	96.0	429	1 APER_PSEAE	O9hyz3 pseudomonas
40	24	96.0	431	1 SYH_PYPAB	O9uy31 pyrococcus
41	24	96.0	449	1 MURC_HELPJ	O92112 helicobacte
42	24	96.0	449	1 MURC_HELPJ	O25340 helicobacte
43	24	96.0	484	1 PARE_HAEIN	P44963 haemophilus
44	24	96.0	555	1 FPRB_MYCLE	O33064 mycobacteri
45	24	96.0	562	1 SIS2_YEAST	P36024 saccharomyc

ALIGNMENTS

?

RESULT 1
VNST_RVFPVZ STANDARD; PRT; 265 AA.
AC P21698;
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Nonstructural protein NS-S.
OS Rift valley fever virus (strain ZH-548 M12) (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxId=11369;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-91111992; PubMed-1846496;
RA Giorgi C., Accardi L., Nicoletti L., Gro M.C., Takehara K.,
RA Hilditch C., Morikawa S., Bishop D.H.L.,
RT "Sequences and coding strategies of the S RNAs of Toscana and Rift
RT Valley fever viruses compared to those of Punta Toro, Sicilian
RT Sandfly fever, and Uukuniemi viruses."
RL Virology 180:738-753(1991).
CC -1- SIMILARITY: NS-S FROM PUNTA TORO, RIFT VALLEY FEVER, SANDFLY FEVER
CC SICILIAN, TOSCANA, AND UUKUNIEMI VIRUSES ARE EVOLUTIONARY RELATED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X53771; CAA37788.1; -
CC PIR: C38552; MNVURV.
CC Nonstructural protein.
KW SEQUENCE 265 AA; 29903 MW; EF2608D29CEB6C84 CRC64;
SQ
Query Match 100.0%; Score 25; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVVPN 5
Db 208 AVVPN 212
RESULT 2
ID DRNL_RABIT STANDARD; PRT; 281 AA.
AC O18998;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Decylnonclease I precursor (EC 3.1.21.1) (DNase I).
GN DNASE1 OR DNLI.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.; SEQUENCE OF 22-41, AND CHARACTERIZATION.
 RC STRAIN-Japanese white; TISSUE-Pancreas, and Urine;
 RX MEDLINE-97361938; PubMed-9230129;
 RA Yasuda T., Takeshita H., Nakajima T., Hosomi O., Nakashima Y.,
 RA Kishi K.;
 RT "Rabbit Dnae I: purification from urine, immunological and
 RT proteochemical characterization, nucleotide sequence, expression in
 RT tissues, relationships with other mammalian Dnae I and phylogenetic
 RT analysis.";
 RL Biochem. J. 325:465-473(1997).
 CC -1- FUNCTION: AMONG OTHER FUNCTIONS, SEEMS TO BE INVOLVED IN CELL
 CC DEATH BY APOPTOSIS. BINDS SPECIFICALLY TO G-ACTIN AND BLOCKS ACTIN
 CC POLYMERIZATION. PREFERENTIALLY ATTACKS DOUBLE-STRANDED DNA AND
 CC PRODUCES OLIGONUCLEOTIDES WITH 5'-PHOSPHO AND 3'-HYDROXY TERMINI.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphodinucleotide and 5'-phosphooligonucleotide end-products.
 CC -1- COFACTOR: DIVALENT CATIONS, PARTICULARLY CALCIUM AND MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: SECRETORY PROTEIN, STORED IN ZYMOGEN
 CC GRANULES AND FOUND IN THE NUCLEAR ENVELOPE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EQUIVALENT LEVELS IN PANCREAS AND PAROTID
 CC GLAND, LOW AMOUNTS IN KIDNEY, LIVER, SMALL INTESTINE, STOMACH AND
 CC THYMUS.
 CC -1- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D82875; BAA21724.1; -;
 DR HSSP: P00639; 2DNU.
 DR InterPro: IPR001582; Dnae_I.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR PRINTS: PR00130; DNASEI.
 DR PRODOM: PD005408; Dnae_I.N; 1.
 DR SMART: SM00476; DNaseIc.1.
 DR PROSITE: PS00919; DNASE_I.1; 1.
 DR PROSITE: PS00918; DNASE_I.2; FALSE_NEG.
 DR KMW Hydroxylase; Endonuclease; Nuclease; Glycoprotein; Calcium; Signal;
 KW Apoptosis; Actin-binding.
 FT SIGNAL 1 21
 FT CHAIN 22 281 DEOXYRIBONUCLEASE I.
 FT DISULFID 122 125 BY SIMILARITY.
 FT ACT_SITE 99 99 BY SIMILARITY.
 FT ACT_SITE 155 155 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .).
 SO SEQUENCE 281 AA; 31345 MW; 6B6535FCB1FE29C64;
 Query Match 100.0%; Score 25; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GN DNASEI OR DNL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE-Heart;
 RX MEDLINE-95160718; PubMed-7857306;
 RA Peitsch M.C., Irmeler M., French L.E., Tschopp J.;
 RT "Genomic organisation and expression of mouse deoxyribonuclease I.";
 RL Biochem. Biophys. Res. Commun. 207:62-68(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE-Kidney, and Parotid gland;
 RX MEDLINE-97335420; PubMed-9192086;
 RA Takeshita H., Yasuda T., Nakajima T., Hosomi O., Nakashima Y.,
 RA Kishi K.;
 RT "Mouse deoxyribonuclease I (Dnae I): biochemical and immunological
 RT characterization, cDNA structure and tissue distribution.";
 RL Biochem. Mol. Biol. Int. 42:65-75(1997).
 CC -1- FUNCTION: AMONG OTHER FUNCTIONS, SEEMS TO BE INVOLVED IN CELL
 CC DEATH BY APOPTOSIS. BINDS SPECIFICALLY TO G-ACTIN AND BLOCKS
 CC ACTIN POLYMERIZATION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphodinucleotide and 5'-phosphooligonucleotide end-products.
 CC -1- COFACTOR: DIVALENT CATIONS, PARTICULARLY CALCIUM AND MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: SECRETORY PROTEIN, STORED IN ZYMOGEN
 CC GRANULES AND FOUND IN THE NUCLEAR ENVELOPE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN STRIATED MUSCLE, KIDNEY, INTESTINE, LIVER,
 CC LYMPHOES, BUT NOT IN THE HEART, SPLEEN OR PANCREAS.
 CC -1- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U00478; AAA03710.1; -;
 DR EMBL: D83038; BAA28622.1; -;
 DR PIR: JC2526; JC2526.
 DR HSSP: P00639; 3DNI.
 DR MCD: MGI103157; DnaeI.
 DR InterPro: IPR001582; Dnae_I.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR PRINTS: PR00130; DNASEI.
 DR PRODOM: PD005408; Dnae_I.N; 1.
 DR SMART: SM00476; DNaseIc.1.
 DR PROSITE: PS00918; DNASE_I.2; 1.
 DR PROSITE: PS00919; DNASE_I.1; 1.
 DR KMW Hydroxylase; Endonuclease; Nuclease; Glycoprotein; Calcium; Signal;
 KW Apoptosis; Actin-binding.
 FT SIGNAL 1 22
 FT CHAIN 23 284 DEOXYRIBONUCLEASE I.
 FT DISULFID 123 126 BY SIMILARITY.
 FT DISULFID 195 231 ESSENTIAL FOR ENZYMATIC ACTIVITY
 FT ACT_SITE 100 100 BY SIMILARITY.
 FT ACT_SITE 156 156 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 239 240 AG -> VR (IN REF. 1).
 SO SEQUENCE 284 AA; 32027 MW; 8682E20515EBB510 CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 284;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 246 AVVFN 250

RESULT 4
CC2A_ANTMA STANDARD: PRT: 294 AA.

AC 038772;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cell division control protein 2 homolog A (EC 2.7.1.1).
GN CDC2A.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Antirrhinaceae; Antirrhineae;
OC Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96434535; PubMed=8837502;
RA Robert P.R., Gaudin V., Lunness P., Coen E.S., Doonan J.H.;
RT "Distinct classes of cdc2-related genes are differentially expressed
during the cell division cycle in plants."
RL Plant Cell 8:1465-1476(1996).
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CELL CYCLE.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: X97637; CAA66233.1; ALT_INIT.
DR HSSP: P24941; IHCL.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF000069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 294 AA: 33912 MW: 0DD450ECDEFA166C CRC64;

Query Match 100.0%; Score 25; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVFN 5
DB 251 AVVFN 255

RESULT 5

CCD2_CHERU STANDARD: PRT: 294 AA.

AC P93101;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein-2 homolog (EC 2.7.1.1.) (p34cdc2).
GN CDC2 OR CDK34.
OS Chenopodium rubrum (Red goosefoot) (Pigweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Chenopodiaceae; Chenopodium.
OX NCBI_TaxID=3560;
RN [1]
RP SEQUENCE FROM N.A.
RA Renz A., Schmelz B., Beck E.;
RT "Nucleotide sequence of a cDNA encoding a CDK34-protein kinase from a
photoautotrophic cell suspension culture of Chenopodium rubrum L.";
RL (in) Plant Gene Register PGR97-045.
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
THE REPLICATIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: Y10160; CAA71242.1;
DR HSSP: P24941; IHCL.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF000069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 294 AA: 33832 MW: B82748FE84B600E6 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVFN 5
DB 251 AVVFN 255

RESULT 6

DAPL_ARATH STANDARD: PRT: 364 AA.

AC 09LZX6; 049355; 09SWS8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Dihydrodipicolinate synthase 1, chloroplast precursor (EC 4.2.1.52)
 DE (DHDPs 1).
 GN DHDP1 OR DHPS1 OR AT3G60880 OR T4C21_290.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eucosids II: Brassicales: Brassicaceae: Arabidopsis.
 OC NCBI_TaxID=3702.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94325478; PubMed=8049377;
 RA Vauterin M., Jacobs M.,
 RT "Isolation of a poplar and an Arabidopsis thaliana dihydrodipicolinate
 RT synthase cDNA clone."
 RL Plant Mol. Biol. 25:545-550(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unselid M.,
 RA Faltmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Delsen M., Boutry M., Grievall L.A., Mache R., Pulgomech P.,
 RA De Simone V., Choise N., Artiguenave F., Robert C., Brothier P.,
 RA Mincker P., Catolico L., Weissenbach J., Saurin W., Queller F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Schaefer M., Schoen O., Bagnues M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,
 RA Cooke R., Landie M., Berger-Llauro C., Purnelle B., Masny D.,
 RA De Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Meyer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Wals A., Utterback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura M., Matsumoto M., Kawashima K., Kishida Y.,
 RA Kiyokawa S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RT thaliana."
 RL Nature 408:820-822(2000).
 RN [3]
 RP SEQUENCE OF 1-158 FROM N.A.
 RA Vauterin M.,
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: L-aspartate 4-semialdehyde + pyruvate =
 CC dihydrodipicolinate + 2 H(2)O.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 CC semialdehyde: first step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the DHDPs family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X72971; CAB45642.1;
 DR EMBL: AL162295; CAB82692.1;
 DR EMBL: X98080; CAB6703.1;
 DR PIR: S46304; S46304.
 DR PIR: T47899; T47899.
 DR HSSP: P05640; IDHP.

DR InterPro: IPR005263; DAP2.
 DR InterPro: IPR002220; DHDPs.
 DR Pfam: PF00701; DHDPs_1.
 DR PRINTS: PR00146; DHPICNSHASE.
 DR ProDom: PD001859; DHDPs_1.
 DR TIGRFAMs: TIGR00674; daps_1.
 DR PROSITE: PS00665; DHDPs_1; 1.
 DR PROSITE: PS00666; DHDPs_2; 1.
 KW Lyase; Diaminopimelate biosynthesis; lysine biosynthesis;
 KW Lyase; Diaminopimelate biosynthesis; transit peptide; Multigene family.
 FT TRANSIT 1 38 CHLOROPLAST (POTENTIAL).
 FT CHAIN 39 364 DIHYDRODIPICOLINATE SYNTHASE 1.
 FT ACT_SITE 221 221 BY SIMILARITY.
 FT CONFLICT 285 285 A->E (IN REF. 1).
 SQ SEQUENCE 364 AA; 40464 MW; 73893B17AA/9E6EE CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 364;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAVPN 5
 Db 39 AAVPN 43
 DAP2_ARATH STANDARD; PRT; 365 AA.
 ID DAP2_ARATH
 AC G9FV8; 022129;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dihydrodipicolinate synthase 2, chloroplast precursor (EC 4.2.1.52)
 DE (DHDPs 2).
 GN DHDP2 OR AT2G45440 OR F4L23.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eucosids II: Brassicales: Brassicaceae: Arabidopsis.
 OC NCBI_TaxID=3702.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Wassilewskaja;
 RA Sarrobert C., Ribaud M.-C., Conlard-David P., Gineste S.,
 RA Bechtold N., Robaglia C., Nussanne L.,
 RT "Accumulation of threonine resulting from mutation in a new
 RT dihydrodipicolinate synthase gene altered Arabidopsis growth."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.,
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 CC -1- CATALYTIC ACTIVITY: L-aspartate 4-semialdehyde + pyruvate =
 CC dihydrodipicolinate + 2 H(2)O.
 CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
 CC semialdehyde: first step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the DHDPs family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF200325; AAG28565.1; -
DR EMBL: AC002387; AAB82620.1; -
DR PIR: E84890; E84890.
DR HSSP: P05640; LDHP.
DR InterPro: IPR005263; DAPA.
DR InterPro: IPR002220; DHPS.
DR Pfam: PF00701; DHPS; 1.
DR PRINTS: PR00146; DHPICSYNTASE.
DR Prodom: PD001859; DHPS; 1.
DR TIGRFAMS: TIGR00674; dapa; 1.
DR PROSITE: PS00665; DHPS_1; 1.
DR PROSITE: PS00666; DHPS_2; 1.
DR Lyase: Diaminopimelate biosynthesis; Lysine biosynthesis;
KW Allosteric enzyme; Chloroplast; Transil peptide; Multigene family.
FT TRANSIT 1 39 CHLOROPLAST (POTENTIAL).
FT CHAIN 40 365 DIHYDRODIPICOLINATE SYNTHASE 2.
FT ACT_SITE 222 222 BY SIMILARITY.
FT CONFLICT 26 26 S -> G (IN REF. 1).
SQ SEQUENCE 365 AA; 40292 MW; 43BECEPFC3E557A CRC64;

Query Match 100.0%; Score 25; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 83;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 40 AVVPN 44

RESULT 8

TH1L_MYCN STANDARD; PRT; 387 AA.
AC P75228;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable thiamine biosynthesis protein th1l.
GN TH1L OR MPN550 OR MP292.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RA MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RT Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY (BY
CC SIMILARITY).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE TH1L FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE000027; AAB95940.1; -
DR PIR: S73618; S73618.
DR HAMAP: MF_00021; -; 1.
DR InterPro: IPR003720; Th1L.

DR InterPro: IPR004114; THUMP_dom.
DR Pfam: PF02568; Th1L; 1.
DR Pfam: PF02926; THUMP; 1.
DR TIGRFAMS: TIGR00342; TIGR00342; 1.
KW Thiamine biosynthesis; Complete proteome.
SQ SEQUENCE 387 AA; 43341 MW; E418669E7A8F64DF CRC64;

Query Match 100.0%; Score 25; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 306 AVVPN 310

RESULT 9

CNRC_ALCEU STANDARD; PRT; 418 AA.
AC P37974;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Nickel and cobalt resistance protein cnrc precursor.
GN CNRC.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Plasmid pMO128.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH34;
RA MEDLINE=93139046; PubMed=8380802;
RA Liesegang H., Lemke K., Siddiqui R.A., Schlegel H.-G.;
RT "Characterization of the inducible nickel and cobalt resistance
RT determinant cnr from pMO128 of Alcaligenes eutrophus CH34.";
RL J. Bacteriol. 175:767-778(1993).
CC -1- FUNCTION: NICKEL AND COBALT RESISTANCE PROTEINS CNRA, CNRB, CNRC
CC CNRH AND CNRR MAY BE INVOLVED IN THE REGULATION OF CNR.
CC -1- SIMILARITY: HIGH, TO A.XYLOSOXYDANS NCCC, AND SOME, TO A.EUTROPHUS
CC CZCC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M91650; AAA21968.1; -
DR PIR: E47056; E47056.
DR InterPro: IPR003423; OEP.
DR Pfam: PF02321; OEP; 2.
KW Plasmid; Transport; Nickel; Cobalt; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 418 NICKEL AND COBALT RESISTANCE PROTEIN
FT SIGNAL 30 418 NICKEL AND COBALT RESISTANCE PROTEIN
SQ SEQUENCE 418 AA; 43760 MW; C7A277D975CE69E5 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 232 AVVPN 236

RESULT 10
SYN_VIBCH STANDARD; PRT; 466 AA.
ID SYN_VIBCH

```

AC O9KSF9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Asparaginyl-tRNA synthetase (EC 6.1.1.22) (Asparagine--tRNA ligase)
DE (AsnRS)
GN ASNS OR VC1297.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Haddad J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004209; AAF94456.1; ALT_INIT.
CC DR HSSP; Q52428; 1B8A.
CC DR TIGR; VC1297; -.
CC DR HAMAP; MF_00534; -; 1.
CC DR InterPro: IPR004522; AsnS.
CC DR InterPro: IPR004364; tRNA-synt_2.
CC DR InterPro: IPR002312; tRNA-synt_asp.
CC DR InterPro: IPR006195; tRNA_ligase_II.
CC DR Pfam; PF01336; tRNA-synt_2; 1.
CC DR Pfam; PF01336; tRNA-anti; 1.
CC DR PRINTS; PRO1042; TRNASYNTHASP.
CC DR TIGRFAMs; TIGR00457; asns.1.
CC DR PROSITE; PSS0862; AA_TRNA_LIGASE_II; 1.
CC DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC DE (AsnRS).
CC GN ASNS OR VC1297.
CC OS Vibrrio cholerae.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC OC Vibrionaceae; Vibrrio.
CC OX NCBI_TaxID=666;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=El Tor N16961 / Serotype O1;
CC RX MEDLINE=20406833; PubMed=10952301;
CC RA Haddad J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
CC RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
CC RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
CC RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
CC RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
CC RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
CC RA Fraser C.M.;
CC RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
CC cholerae."
CC RL Nature 406:477-483(2000).
CC CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004209; AAF94456.1; ALT_INIT.
CC DR HSSP; Q52428; 1B8A.
CC DR TIGR; VC1297; -.
CC DR HAMAP; MF_00534; -; 1.
CC DR InterPro: IPR004522; AsnS.
CC DR InterPro: IPR004364; tRNA-synt_2.
CC DR InterPro: IPR002312; tRNA-synt_asp.
CC DR InterPro: IPR006195; tRNA_ligase_II.
CC DR Pfam; PF01336; tRNA-synt_2; 1.
CC DR Pfam; PF01336; tRNA-anti; 1.
CC DR PRINTS; PRO1042; TRNASYNTHASP.
CC DR TIGRFAMs; TIGR00457; asns.1.
CC DR PROSITE; PSS0862; AA_TRNA_LIGASE_II; 1.
CC DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC DE (AsnRS).
CC GN ASNS OR VC1297.
CC OS Vibrrio cholerae.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC OC Vibrionaceae; Vibrrio.
CC OX NCBI_TaxID=666;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=El Tor N16961 / Serotype O1;
CC RX MEDLINE=20406833; PubMed=10952301;
CC RA Haddad J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
CC RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
CC RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
CC RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
CC RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
CC RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
CC RA Fraser C.M.;
CC RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
CC cholerae."
CC RL Nature 406:477-483(2000).

```

```

DE (AsnRS).
GN ASNS OR VP1893.
OS Vibrrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005079; BAC60156.1; -.
CC DR HAMAP; MF_00534; -; 1.
CC DR PROSITE; PSS0862; AA_TRNA_LIGASE_II; 1.
CC DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC DE (AsnRS).
CC GN ASNS OR VVI2254.
CC OS Vibrrio vulnificus.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC OC Vibrionaceae; Vibrrio.
CC OX NCBI_TaxID=672;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=CMC6;
CC RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
CC Choy H.E.;
CC RT "Complete genome sequence of Vibrio vulnificus CMC6."
CC Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC EMBL; AP005079; BAC60156.1; -.
CC DR HAMAP; MF_00534; -; 1.
CC DR PROSITE; PSS0862; AA_TRNA_LIGASE_II; 1.
CC DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC DE (AsnRS).
CC GN ASNS OR VVI2254.
CC OS Vibrrio vulnificus.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC OC Vibrionaceae; Vibrrio.
CC OX NCBI_TaxID=672;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=CMC6;
CC RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
CC Choy H.E.;
CC RT "Complete genome sequence of Vibrio vulnificus CMC6."
CC Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE016804; AAO10633.1; -
 DR HAMM; ME_00534; -; 1.
 DR InterPro: IPR004522; AsnS.
 DR InterPro: IPR004364; tRNA-synt_2.
 DR InterPro: IPR002312; tRNA-synt_2.
 DR InterPro: IPR004365; tRNA-anti.
 DR InterPro: IPR006195; tRNA_ligase_II.
 DR Pfam: PF00152; tRNA-synt_2; 1.
 DR Pfam: PF01336; tRNA-anti; 1.
 DR PRINTS: PR01042; TRNASYNTHASP.
 DR TIGRFAMs: TIGR00457; asns; 1.
 DR PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
 KW Aminocyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 466 AA; 52406 MW; A4577F799612F58B CRC64;

Query Match 100.0%; Score 25; DB 1; Length 466;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVFN 5
 Db 52 AVVFN 56

RESULT 13
 NIFA_AZOI STANDARD; PRT; 624 AA.
 ID P54929;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nif-specific regulatory protein.
 GN NIFA.
 OS Azospirillum lipoferrum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Azospirillum.
 OX NCBI_TaxID=193;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shigematsu T., Hidaka M., Masaki H., Uozumi T.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NIFA, A TRANSCRIPTIONAL ACTIVATOR, IS REQUIRED FOR
 CC ACTIVATION OF MOST NIF OPERONS, WHICH ARE DIRECTLY INVOLVED IN
 CC NITROGEN FIXATION. NIFA INTERACTS WITH SIGMA-54.
 CC -1- SIMILARITY: Contains 1 sigma-54 factor interaction ATP-binding
 CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D13799; BAA02956.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR002197; HTH_Fis.
 DR InterPro: IPR002078; Sig54_interact.
 DR Pfam: PF01590; GAF; 1.
 DR Pfam: PF02954; HTH_8; 1.
 DR Pfam: PF00158; Sigma54_activat; 1.
 DR PRINTS: PR01590; HTHFIS.

DR SMART; SM00382; AAA; 1.
 DR SMART; SM00065; GAF; 1.
 DR TIGRFAMs: TIGR01199; HTH_fis; 1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE: PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE: PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE: PS50045; SIGMA54_INTERACT_4; 1.
 KW Nitrogen fixation; Transcription regulation; Activator;
 KW ATP-binding; DNA-binding; Metal-binding;
 FT DOMAIN 203 431 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 FT 432 581 INTER-DOMAIN LINKER.
 FT DOMAIN 582 624 C-TERMINAL DNA-BINDING DOMAIN.
 FT NP_BIND 231 238 ATP (POTENTIAL).
 FT METAL 294 303 ATP (POTENTIAL).
 FT METAL 445 445 BY SIMILARITY.
 FT METAL 450 450 BY SIMILARITY.
 FT DNA_BIND 596 615 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 624 AA; 67119 MW; 223A56AD444E7016 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 624;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVFN 5
 Db 96 AVVFN 100

RESULT 14
 NIFA_AZOBR STANDARD; PRT; 625 AA.
 ID P30667;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nif-specific regulatory protein.
 GN NIFA.
 OS Azospirillum brasilense.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Azospirillum.
 OX NCBI_TaxID=192;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sp7;
 RX MEDLINE=92140038; PubMed=1779763;
 RA Liang Y.Y., Kaminski P.A., Elmerich C.;
 RT "Identification of a nifA-like regulatory gene of Azospirillum
 RT brasilense Sp7 expressed under conditions of nitrogen fixation and in
 RT the presence of air and ammonia.";
 RL Mol. Microbiol. 5:2735-2744(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=93122522; PubMed=1362170;
 RA Liang Y.Y., de Zamaroczy M., Arsene F., Paquelin A., Elmerich C.;
 RT "Regulation of nitrogen fixation in Azospirillum brasilense Sp7:
 RT involvement of nifA, glnA and glnB gene products.";
 RL FEMS Microbiol. Lett. 79:113-119(1992).
 CC -1- FUNCTION: NIFA, A TRANSCRIPTIONAL ACTIVATOR, IS REQUIRED FOR
 CC ACTIVATION OF MOST NIF OPERONS, WHICH ARE DIRECTLY INVOLVED IN
 CC NITROGEN FIXATION. NIFA INTERACTS WITH SIGMA-54.
 CC -1- SIMILARITY: Contains 1 sigma-54 factor interaction ATP-binding
 CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X60714; CAA43126.1; -
 DR PIR: S18420; S18420.

DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR002197; HTH_Fls.
 DR InterPro: IPR002078; Sig54_interact.
 DR Pfam: PF01590; GAF.1.
 DR Pfam: PF02954; HTH_8.1.
 DR Pfam: PF00158; Sigma54_activat.1.
 DR PRINTS: PR01590; HTHFIS.
 DR SMART: SM00382; AAA.1.
 DR SMART: SM00065; GAF.1.
 DR TIGRfams: TIGR01199; HTH_fas.1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1.1.
 DR PROSITE: PS00676; SIGMA54_INTERACT_2.1.
 DR PROSITE: PS00688; SIGMA54_INTERACT_3.1.
 DR PROSITE: PS50045; SIGMA54_INTERACT_4.1.
 DR Nitrogen fixation; Transcription regulation; Activator;
 KW ATP-binding; DNA-binding.
 KM DOMAIN 205 433 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 FT DOMAIN 434 582 INTER-DOMAIN LINKER.
 FT DOMAIN 583 625 C-TERMINAL DNA-BINDING DOMAIN.
 FT NP_BIND 233 240 ATP (POTENTIAL).
 FT NP_BIND 296 305 ATP (POTENTIAL).
 FT METAL 447 447 BY SIMILARITY.
 FT METAL 452 452 BY SIMILARITY.
 FT DNA_BIND 597 616 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 625 AA; 67855 MW; D060EA86CB1FECC3 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 625;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 96 AVPN 100

RESULT 15

FLXY_CAUCR STANDARD: PRT; 954 AA.
 AC P15345; P15346;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Regulatory protein flaxy.
 GN FLAXY OR FLAX/FLAY OR CCI465.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP PRELIMINARY SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=89178645; PubMed=2648000;
 RA Kaplan J.B., Dingwall A., Bryan R., Chamber R., Shapiro L.;
 RT "Temporal regulation and overlap organization of two Caulobacter
 flagellar genes.";
 RL J. Mol. Biol. 205:71-83(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Kutterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- FUNCTION: FUNCTIONS IN TRANS TO MODULATE THE LEVEL OF
 CC TRANSCRIPTION OF THE FLAGELLIN GENES AND SEVERAL GENES ENCODING
 CC CHEMOTAXIS FUNCTIONS. IT IS ITSELF TEMPORALLY CONTROLLED.

CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS. IT ENCODES FOR TWO SEPARATE ADJACENT ORFS, FLAX AND
 CC FLAY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AE005820; AAK23445.1; -
 DR PIR: A87431; A87431.
 DR TIGR: CCI465; -
 KW Flagella; Transcription regulation; Trans-acting factor;
 CC Complete proteome.
 SQ SEQUENCE 954 AA; 97979 MW; BC565B4E8FB6805D CRC64;

Query Match 100.0%; Score 25; DB 1; Length 954;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPN 5
 DB 36 AVPN 40

Search completed: September 4, 2003, 21:01:43
 Job time : 4.33824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 : Search time 20.5147 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893d-58
Perfect score: 25
Sequence: 1 AVVPN 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525.seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._viral:*
16: sp._bacteriap:*
17: sp._archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	134	2	Q53805 streptomyce
2	25	100.0	161	16	Q92X31 rhizobium m
3	25	100.0	174	2	Q9ZG06 staphylococ
4	25	100.0	200	12	073538 rift valley
5	25	100.0	200	12	073544 rift valley
6	25	100.0	200	12	073540 rift valley
7	25	100.0	200	12	083JF9 rift valley
8	25	100.0	200	12	073541 rift valley
9	25	100.0	200	12	073543 rift valley
10	25	100.0	200	12	073548 rift valley
11	25	100.0	200	12	073537 rift valley
12	25	100.0	200	12	073547 rift valley
13	25	100.0	200	12	072702 bungavirus
14	25	100.0	200	12	073542 rift valley
15	25	100.0	200	12	073539 rift valley
16	25	100.0	200	12	072701 lunyo virus

17	25	100.0	200	12	073545 rift valley
18	25	100.0	200	12	073549 rift valley
19	25	100.0	200	12	073535 rift valley
20	25	100.0	200	12	073536 rift valley
21	25	100.0	200	12	073546 rift valley
22	25	100.0	217	10	092R25 populus tre
23	25	100.0	217	17	08TY62 methanopyru
24	25	100.0	229	17	09UZX5 pyrococcus
25	25	100.0	234	10	08S7U8 oryza sativ
26	25	100.0	242	4	09H1G7 homo sapien
27	25	100.0	247	5	08IR10 drosophila
28	25	100.0	284	11	091WL7 mus musculus
29	25	100.0	292	16	0929E2 listeria in
30	25	100.0	292	16	08Y433 listeria mo
31	25	100.0	296	2	086189 erwinia chr
32	25	100.0	301	2	088189 enterococu
33	25	100.0	312	10	09L079 oryza sativ
34	25	100.0	326	10	09S0K8 lycopersico
35	25	100.0	332	11	08K0U1 mus musculu
36	25	100.0	334	16	08YS25 anabaena sp
37	25	100.0	334	16	09CFV7 lactococcus
38	25	100.0	348	16	08C2X4 yersinia pe
39	25	100.0	351	16	09PF30 xanthomonas
40	25	100.0	351	16	08P125 xanthomonas
41	25	100.0	353	16	08P6S6 aeropyrum p
42	25	100.0	363	17	09YCS0 arabidopsis
43	25	100.0	365	10	08VZQ1 brugia mala
44	25	100.0	381	5	09GT27 hydrangea m
45	25	100.0	386	12	091MN4

ALIGNMENTS

RESULT 1
ID Q53805 PRELIMINARY; PRT; 134 AA.
AC 053805;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Pre-phenomycin precursor.
GN PHM.
OS Streptomyces roseoveriticillatus.
OC Bacteria: Actinobacteria: Actinobacteriales:
OC Streptomycineae: Streptomycetaceae: Streptomycetes.
OX NCBI_TaxID=66429;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC29809;
RX MEDLINE=92104838; PubMed=1761418;
RA Muramatsu R., Abe S., Hayashi H., Yamaguchi K., Jinda K., Sakano K.,
RA Inouye Y., Nakamura S.;
RT "Complete amino acid sequence of phenomycin, an antitumor polypeptide
antibiotic."
RL J. Antibiot. 44:1222-1227(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC29809;
RX MEDLINE=94230131; PubMed=8175490;
RA Sakata N., Oka T., Ikono S., Hori M.;
RT "Nucleotide sequence of the phenomycin gene from Streptoveriticillium
baldaei Ma564-C1".
RL J. Antibiot. 47:370-371(1994).
DR EMBL; D17759; BAA04606.1; -;
KW Signal.
FT CHAIN 1 45 POTENTIAL.
FT SIGNAL 46 134 PHENOMYCIN.
SQ SEQUENCE 134 AA; 13751 MW; 577ADD7C503418E CRC64;
Query Match 100.0%; Score 25; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 DB 42 AVVPN 46

RESULT 2

Q92X31 PRELIMINARY; PRT; 161 AA.
 AC Q92X31;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative oxidoreductase protein.
 GN RB0131 OR SMe20131.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymB (megaplasmid 2).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P., Guzy J.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Guzy J.,
 RA Golding B., Fuenler A.;
 RT "The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603642; CAC48531.1; -
 DR InterPro; IPR002888; 2Fe-2S bind.
 DR InterPro; IPR006508; 2Fe-2S_Ferredoxin.
 DR InterPro; IPR001041; Ferredoxin.
 DR Pfam; PF00111; fer2_1.
 DR Pfam; PF01799; fer2_2; 1.
 DR PRODOM; PD186071; 2Fe-2S_bind; 1.
 DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
 KW Plasmid; hypothetical protein; Complete proteome.
 SQ SEQUENCE 161 AA; 1782 MW; 7128EAFCFBFC1F7 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 161;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 DB 17 AVVPN 21

RESULT 3

Q9ZG06 PRELIMINARY; PRT; 174 AA.
 AC Q9ZG06;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Type-I signal peptidase SipA.
 GN SIPA.
 OS Staphylococcus carnosus.
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1281;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TM300;
 RA Matzen A., Frendl R.;
 RT "Cloning and characterization of the type-I signal peptidase of
 RT Staphylococcus carnosus.";
 RT Submitted (SFP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF089862; AAD09010.1; -
 DR HSSP; P00803; 1B12.
 DR InterPro; IPR000508; SigPase.
 DR Pfam; PF00461; Peptidase_S26; 1.

SQ SEQUENCE 174 AA; 19931 MW; 965B7A8FDB818C2B CRC64;

Query Match 100.0%; Score 25; DB 2; Length 174;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 DB 29 AVVPN 33

RESULT 4

Q73538 PRELIMINARY; PRT; 200 AA.
 AC Q73538;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NSs protein (Fragment).
 GN NSs.
 OS Rift valley fever virus (RVFV).
 CC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 OX NCBI_TaxID=11588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ar D104769;
 RX MEDLINE=98033195; PubMed=9367372;
 RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
 RA Thiongane Y., Bouloy M.;
 RT "Variability of the NSs protein among rift valley fever virus
 RT isolates.";
 RL J. Gen. Virol. 78:2853-2858(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ar D104769;
 RA Prehaud C.J.N.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y12742; CAA73276.1; -
 FT NON_TER 1 1
 FT NON_TER 200 200
 SQ SEQUENCE 200 AA; 22558 MW; 1DF3740A939F36GD CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
 DB 187 AVVPN 191

RESULT 5

Q73544 PRELIMINARY; PRT; 200 AA.
 AC Q73544;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NSs protein (Fragment).
 GN NSs.
 OS Rift valley fever virus (RVFV).
 CC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 OX NCBI_TaxID=11588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H D47311;
 RX MEDLINE=98033195; PubMed=9367372;
 RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
 RA Thiongane Y., Bouloy M.;
 RT "Variability of the NSs protein among rift valley fever virus
 RT isolates.";
 RL J. Gen. Virol. 78:2853-2858(1997).
 RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-H DA7311;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12748; CAA73282.1; -
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22572 MW; F3F3740A939FC68 CRC64;

Query Match
Best Local Similarity 100.0%; Score 25; DB 12; Length 200;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 6
ID 073540 PRELIMINARY; PRT; 200 AA.
AC 073540;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-Ar D38661;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT Variability of the NSS protein among rift valley fever virus
RT isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar D38661;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12744; CAA73278.1; -
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22544 MW; 485876F739FDA45F CRC64;

Query Match
Best Local Similarity 100.0%; Score 25; DB 12; Length 200;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 7
ID 08JUF9 PRELIMINARY; PRT; 200 AA.
AC 08JUF9;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA01-1322;

RA Miller B.R., Godsey M.S., Crabtree M.B., Al-Mazrao Y., Al-Jeffri M.H.,
RA Abdoon A.M., Al-Seghayer S.M., Ksiazek T.G.;
RT "Out of Africa: Isolation and Genetic Characterization of Rift Valley
RT Fever Virus from Aedes arabiensis from the Asir Region of the Kingdom
RT of Saudi Arabia."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF393744; AA073694.1; -
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22530 MW; 217BE1B99C34C6B6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 25; DB 12; Length 200;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 8
ID 073541 PRELIMINARY; PRT; 200 AA.
AC 073541;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-Ar M8811;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT Variability of the NSS protein among rift valley fever virus
RT isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar M8811;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12745; CAA73279.1; -
FT NON_TER 1
FT NON_TER 200
SQ SEQUENCE 200 AA; 22502 MW; EA7EFB79D8B3419 CRC64;

Query Match
Best Local Similarity 100.0%; Score 25; DB 12; Length 200;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 9
ID 073543 PRELIMINARY; PRT; 200 AA.
AC 073543;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-B Eg93;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-B Eg93;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12747; CAA73281.1; -
FT NON_TER 1 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 22592 MW; C7C76FEEBB3C98DF CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 10
OY 073548 PRELIMINARY; PRT; 200 AA.
ID 073548;
AC 073548;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MP12;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MP12;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12754; CAA73288.1; -
FT NON_TER 1 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 22564 MW; D8C87FEFEBB2B8DFD CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 11
OY 073537 PRELIMINARY; PRT; 200 AA.
ID 073537;
AC 073537;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar B1976;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Ar B1976;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12741; CAA73275.1; -
FT NON_TER 1 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 22533 MW; 271D6D130A4F7E9B CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 12
OY 073547 PRELIMINARY; PRT; 200 AA.
ID 073547;
AC 073547;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NSS protein (Fragment).
GN NSS.
OS Rift valley fever virus (RVFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H Eg93A;
RX MEDLINE=98033195; PubMed=9367372;
RA Sall A.A., de A Zanotto P.M., Zeller H.G., Digoutte J.P.,
RA Thiongane Y., Bouloy M.;
RT "Variability of the NSs protein among rift valley fever virus
isolates.";
RL J. Gen. Virol. 78:2853-2858(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H Eg93A;
RA Prehaud C.J.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y12751; CAA73285.1; -
FT NON_TER 1 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 22654 MW; 823F356F756E3AF3 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVVPN 5
DB 187 AVVPN 191

RESULT 13

072702

PRELIMINARY; PRT; 200 AA.

AC 072702; 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

NS protein (Fragment).

GN NSS.

OS Bunyavirus sp. 'strain SNS'.

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.

OX NCBI_TaxID=75386;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SNS;

RX MEDLINE-98033195; PubMed-9367372;

RA Sall A.A., de A Zanolto P.M., Zeller H.G., Digoutte J.P.,

Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus

isolates".

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-SNS;

RX Prehaud C.J.N.;

RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL; Y12755; CAA73289.1; -.

FT NON_TER

FT NON_TER

SQ SEQUENCE

200 AA; 22472 MW; E7551856FE61CD88 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5

DB 187 AVVPN 191

RESULT 14

073542

PRELIMINARY; PRT; 200 AA.

AC 073542;

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

NSS protein (Fragment).

GN NSS.

OS Rift valley fever virus (RVFV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

OX NCBI_TaxID=11588;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar Mg990;

RX MEDLINE-98033195; PubMed-9367372;

RA Sall A.A., de A Zanolto P.M., Zeller H.G., Digoutte J.P.,

Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus

isolates".

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar Mg990;

RA Prehaud C.J.N.;

RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL; Y12746; CAA73280.1; -.

FT NON_TER

FT NON_TER

SQ SEQUENCE

200 AA; 22514 MW; FBA810C78CB64A20 CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5

DB 187 AVVPN 191

RESULT 15

073539

PRELIMINARY; PRT; 200 AA.

AC 073539;

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

NSS protein (Fragment).

GN NSS.

OS Rift valley fever virus (RVFV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

OX NCBI_TaxID=11588;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar D38457;

RX MEDLINE-98033195; PubMed-9367372;

RA Sall A.A., de A Zanolto P.M., Zeller H.G., Digoutte J.P.,

Thiongane Y., Bouloy M.;

RT "Variability of the NSS protein among rift valley fever virus

isolates".

RL J. Gen. Virol. 78:2853-2858(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Ar D38457;

RX Prehaud C.J.N.;

RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL; Y12743; CAA73277.1; -.

FT NON_TER

FT NON_TER

SQ SEQUENCE

200 AA; 22276 MW; 34A2CA0161E8B4DB CRC64;

Query Match 100.0%; Score 25; DB 12; Length 200;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVVPN 5

DB 187 AVVPN 191

Search completed: September 4, 2003, 21:06:40

Job time : 21.6056 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 : Search time 30.8824 Seconds
(Without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893d-59

Perfect score: 31

Sequence: 1 GAVVFN 6

Scoring table: BLOSUM62

Gapop 10.0, Capext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: A_Geneseq_19Jun03.*

- Listing first 45 summaries
- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
 - 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	6	23	AAE16722
2	31	100.0	6	23	AAE16724
3	31	100.0	8	23	AAE16727
4	31	100.0	8	23	AAE16730
5	30	96.8	79	22	AAU58250
6	29	93.5	867	20	AAU34508
7	29	93.5	875	20	AAU34381
8	28	90.3	337	22	ABBS9302
9	28	90.3	428	18	AAW34686

10	28	90.3	428	18	AAW24481
11	28	90.3	428	19	AAW71617
12	28	90.3	428	22	AAE05070
13	28	90.3	469	23	ABBS7384
14	28	90.3	497	22	AAU16361
15	28	90.3	497	24	ABU55430
16	28	90.3	534	22	ABU15908
17	28	90.3	534	24	ABU54977
18	28	90.3	547	23	ABBS92706
19	28	90.3	707	22	AAW33718
20	28	90.3	707	22	AAW33234
21	28	90.3	771	20	AAU34574
22	28	90.3	786	20	AAU34431
23	28	90.3	893	24	ABJ19186
24	28	90.3	901	23	ABP40544
25	28	90.3	944	22	AAU03359
26	28	90.3	1938	24	ABBS8388
27	27	87.1	75	24	ABP75905
28	27	87.1	89	22	AAU59414
29	27	87.1	101	23	ABBS6357
30	27	87.1	101	22	AAU45639
31	27	87.1	103	22	ABBS6076
32	27	87.1	103	22	AAW5385
33	27	87.1	208	23	ABP25631
34	27	87.1	231	21	AAU46950
35	27	87.1	247	22	AAU46950
36	27	87.1	252	20	AAU37064
37	27	87.1	271	24	ABP77868
38	27	87.1	359	23	ABBS3778
39	27	87.1	375	22	ABG26783
40	27	87.1	394	22	ABG26781
41	27	87.1	395	21	AAU46949
42	27	87.1	442	23	ABBS6130
43	27	87.1	637	21	ABBS2694
44	27	87.1	635	23	ABP40686
45	27	87.1	663	22	ABBS9368

ALIGNMENTS

RESULT 1	AAE16722	standard; peptide: 6 AA.
XX	AAE16722	
AC	AAE16722	
XX	09-APR-2002	(first entry)
DT	XX	
XX	XX	
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #2.	
XX	XX	
KW	Hydrolysed elastin peptide; HEP, therapy; cosmetic appearance; cardiant;	
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;	
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;	
KW	hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;	
KW	restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.	
XX	XX	
OS	Ligamentum nuchae.	
OS	Synthetic.	
XX	WO200191700-A2.	
PN	06-DEC-2001.	
PD	XX	
XX	30-MAY-2001; 2001WO-US17384.	
FE	30-MAY-2000; 2000US-0580110.	
XX	PR 30-MAY-2000; 2000US-0580156.	
PR	30-MAY-2000; 2000US-0580893.	
PR	30-MAY-2000; 2000US-0584001.	
XX	XX	
XX	XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGING LLC.	
XX	XX	

Class II EPSs use
Bacillus subtilis
Bacillus subtilis
Novel human protei
Human novel secret
Human novel polype
Human novel polype
Human novel polype
Herbically activ
Human polypeptide,
Human polypeptide sequ
Porphyromonas ging
Porphyromonas ging
Pathogen specific
Staphylococcus epi
P. patens active p
Streptomyces virid
Human secretory po
Protonibacterium
Human zinc finger
Protonibacterium
Human testicular a
Human reproductive
Streptococcus poly
Arabidopsis thalia
C glutamicum prote
Protein involved i
N. gonorrhoeae aml
Lactococcus lactis
Novel human diagno
Novel human diagno
Arabidopsis thalia
Zygascus hellenic
Soybean auxin tran
Staphylococcus epi
Drosophila melanog

PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 6 AA:
XX
Query Match 100.0%; Score 31; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAVVFN 6
IIIIII
DB 1 GAVVFN 6
RESULT 2
AAE16724
ID AAE16724 standard; peptide: 6 AA.
XX
AC AAE16724;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #4.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT Modified-site 6 /note="C-terminal amide"
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGINERLING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;

XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 6 AA:
XX
Query Match 100.0%; Score 31; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAVVFN 6
IIIIII
DB 1 GAVVFN 6
RESULT 3
AAE16727
ID AAE16727 standard; peptide: 8 AA.
XX
AC AAE16727;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
XX
OS Ligamentum nuchae.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT Modified-site 6 /note="C-terminal amide"
XX
PN WO200191700-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.
PR 30-MAY-2000; 2000US-0584001.
XX
PA (CONN-) CONNECTIVE TISSUE IMAGINERLING LLC.
XX
PI Mlts TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin

XX Claim 24; Page 21; 53pp; English.
 PS
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 SQ
 OY Sequence 8 AA;
 Query Match 100.0%; Score 31; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 GAVVPPN 6
 111111
 2 GAVVPPN 7
 RESULT 4
 AAE16730
 ID AAE16730 standard; peptide; 8 AA.
 AC AAE16730;
 DT 09-APR-2002 (first entry)
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; arteriosclerotic;
 KW restenosis post-angioplasty; antitanginal; anticoagulant; thrombolytic;
 KW cyclic.
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"
 FT Modified-site 8
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 XX
 PN WO200191700-A2.
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING L.L.C.
 XX
 PI Mites TF, Sandberg LB, Jimenez F;
 XX

DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 PS
 CC Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived cyclic peptide.
 SQ
 OY Sequence 8 AA;
 Query Match 100.0%; Score 31; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 GAVVPPN 6
 111111
 2 GAVVPPN 7
 RESULT 5
 AAU58250
 ID AAU58250 standard; Protein; 79 AA.
 AC AAU58250;
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #19146.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 OS
 XX
 XX Propionibacterium acnes.
 OS
 PN WO200181581-A2.
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59589.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris
 XX

PS Example 1; SEQ ID No 19445; 1069bp; English.

XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 79 AA:

Query Match 96.8%; Score 30; DB 22; Length 79;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVNP 6
DB 40 GAVIPN 45

RESULT 6
AAV34508

ID AAV34508 standard; Protein: 867 AA.

XX AAV34508;

XX 20-MAR-2003 (updated)

DT 25-AUG-1999 (first entry)

DE Porphyromonas gingivalis protein PG47.

XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis;

KW vaccine; antigenic.

XX Porphyromonas gingivalis.

OS WO9929870-A1.

PD 17-JUN-1999.

PF 10-DEC-1998; 98WO-AU01023.

PR 10-DEC-1997; 97AU-0000839.

PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001546.

PR 10-MAR-1998; 98AU-0002264.

PR 09-APR-1998; 98AU-0002911.

PR 23-APR-1998; 98AU-0003128.

PR 05-MAY-1998; 98AU-0003338.

PR 22-MAY-1998; 98AU-0003654.

PR 29-JUL-1998; 98AU-0004917.

PR 30-JUL-1998; 98AU-0004963.

PR 04-AUG-1998; 98AU-0005028.

PA (CSLC-) CSL LTD.

PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

PI Ross BC, Rothel LJ, Webb EA;

XX

DR WPI; 1999-385613/32.

DR N-PSDB; AAX91726.

PT Antigenic Porphyromonas gingivalis peptides for preventing

PT gingivitis

XX Claim 1; Page 491-492; 588bp; English.

PS AAX91536 to AAX91801 encode two hundred and sixty six antigenic

CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAV34318 to

CC AAV34583. AAX91802 to AAX91989 represent PCR primers used in the

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial

CC activity with a vaccine mechanism of action. The PG polypeptides can be

CC used as vaccines especially against Porphyromonas gingivalis. Probes can

CC be used to detect Porphyromonas gingivalis in standard hybridisation

CC assays. Porphyromonas gingivalis is involved in periodontal disease

CC especially gingivitis.

CC (Updated on 20-MAR-2003 to correct PR field.)

XX Sequence 867 AA;

Query Match 93.5%; Score 29; DB 20; Length 867;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVNP 6
DB 619 GAVIPN 624

RESULT 7
AAV34381

ID AAV34381 standard; Protein: 875 AA.

XX AAV34381;

XX 20-MAR-2003 (updated)

DT 25-AUG-1999 (first entry)

DE Porphyromonas gingivalis protein PG47.

XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis;

KW vaccine; antigenic.

XX Porphyromonas gingivalis.

OS WO9929870-A1.

PD 17-JUN-1999.

PF 10-DEC-1998; 98WO-AU01023.

PR 10-DEC-1997; 97AU-0000839.

PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001546.

PR 10-MAR-1998; 98AU-0002264.

PR 09-APR-1998; 98AU-0002911.

PR 23-APR-1998; 98AU-0003128.

PR 05-MAY-1998; 98AU-0003338.

PR 22-MAY-1998; 98AU-0003654.

PR 29-JUL-1998; 98AU-0004917.

PR 30-JUL-1998; 98AU-0004963.

PR 04-AUG-1998; 98AU-0005028.

PA (CSLC-) CSL LTD.

PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

PI Ross BC, Rothel LJ, Webb EA;

XX WPI; 1999-385613/32.

XX N-PSDB; AAX91599.

XX Antigenic Porphyromonas gingivalis peptides for preventing

XX

PT gingivitis
 XX
 PS Claim 1: Page 346-348; 588pp; English.
 XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (Pg) polypeptide sequences given in AAX94318 to
 CC AAX94353. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the Pg polypeptides. The Pg polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The Pg polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis.
 CC (updated on 20-MAR-2003 to correct PR field.)
 CC
 XX
 SQ Sequence 875 AA;
 Query Match 93.5%; Score 29; DB 20; Length 875;
 Best Local Similarity 66.7%; Pred. No. 6.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GAVVPN 6
 ||::||
 Db 627 GATIPN 632
 RESULT 8
 ABB59302
 ID ABB59302 standard; Protein; 337 AA.
 XX
 AC ABB59302;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 4698.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PMD, Myers EM;
 XX
 DR WPI: 2001-656660/75.
 DR N-PSDB; ABL03405.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 4698; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB163051), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins
 CC (AB5737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 337 AA;
 Query Match 90.3%; Score 28; DB 22; Length 337;
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GAVVPN 6
 ||::||
 Db 221 GSVVPN 226
 RESULT 9
 AAW34686
 ID AAW34686 standard; Protein; 428 AA.
 XX
 AC AAW34686;
 XX
 DT 17-FEB-1998 (first entry)
 XX
 DE Class II EPSP synthase (EPSPs) from Bacillus subtilis.
 XX
 KM 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; glyphosate;
 KM shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
 KM 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
 KM glyphosate resistance gene; glyphosate-tolerance; promoter.
 XX
 OS Bacillus subtilis.
 PN
 FT Key Location/Qualifiers
 FT FH 190..194
 FT Region /label= "characteristic_region
 FT /note= "see AAW34690"
 FT 17..20
 FT Region /label= "characteristic_region
 FT /note= "see AAW34691"
 FT 164..168
 FT Region /label= "characteristic_region
 FT /note= "see AAW34692"
 FT 257..260
 FT Region /label= "characteristic_region
 FT /note= "see AAW34693"
 FT
 PN US5627061-A.
 XX
 PD 06-MAY-1997.
 XX
 PF 07-JUN-1995; 95US-0476008.
 XX
 PR 13-SEP-1994; 94US-0306063.
 PR 31-AUG-1990; 90US-0576537.
 PR 28-AUG-1991; 91US-0749611.
 XX
 PA (MONS) MONSANTO CO.
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
 XX
 DR WPI: 1997-271315/24.
 DR N-PSDB; AAT93793.
 XX
 PT Production of glyphosate-herbicide tolerant plants - using DNA
 PT encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
 PT enzyme(s)
 XX
 PS Claim 6; Fig 18; 151pp; English.
 XX
 CC AAW34683-89 represent a new class of glyphosate-tolerant
 CC 5-enolpyruvylshikimate-3-phosphate synthases (EPSPs). These novel
 CC EPSPs enzymes have little homology with known Class I EPSPs enzymes, and
 CC belong to a new class, Class II. The present sequence was isolated from
 CC Bacillus subtilis. The EPSPs enzymes are part of the shikimic
 CC acid pathway, which leads to the biosynthesis of aromatic compounds.

CC EPSPS converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid
 CC to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the
 CC herbicide glyphosate. It would be useful to produce transgenic crops
 CC containing glyphosate resistance genes so that glyphosate-containing
 CC herbicides can be applied to selectively kill weeds. The novel EPSPS
 CC enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that
 CC when introduced into a plant, the plant is made glyphosate-tolerant, and
 CC EPSPS enzyme activity is not affected. These class II EPSPS enzymes are
 CC fused to a chloroplast transit peptide to target the protein into the
 CC chloroplast, which is the site for the shikimic acid pathway. In
 CC addition, the EPSPS gene is cloned into a plant under the control of a
 CC promoter such as figwort mosaic virus promoter or the cauliflower mosaic
 CC virus promoter, so that expression is enhanced.

XX Sequence 428 AA;

Query Match 90.3%; Score 28; DB 18; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 Db 241 GAVVFN 246

RESULT 10
 AAW24481
 ID AAW24481 standard; Protein; 428 AA.

AC AAW24481;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-OCT-1997 (first entry)

XX Class II EPSPS used for glyphosate resistant plant production.

DE 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
 KM glyphosate resistant; transgenic plant; herbicide; shikimic acid;
 KM fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
 XX tobacco.

XX Bacillus subtilis.

OS US5633435-A.

PN 27-MAY-1997.

PD 13-SEP-1994; 94US-0306063.

PF 13-SEP-1994; 94US-0306063.

PR 13-SEP-1994; 94US-0306063.

PR 31-AUG-1990; 90US-0576537.

PR 28-AUG-1991; 91US-0749611.

XX (MONS) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

XX WPI; 1997-297418/27.

XX N-PSDB; AAT77317.

XX New isolated 5-enolpyruvylshikimate-3-phosphate synthase gene -
 PT used for transforming plants to produce plants which are tolerant to
 PT glyphosate herbicide

XX Claim 7; Column 97-100; 154pp; English.

CC AAW24481 shows the sequence of a class II 5-enolpyruvylshikimate-3-
 CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
 CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
 CC the production of herbicide resistant (glyphosate-tolerant) plants
 CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
 CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
 CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 428 AA;

Query Match 90.3%; Score 28; DB 18; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
 Db 241 GAVVFN 246

RESULT 11
 AAW71617
 ID AAW71617 standard; Protein; 428 AA.

AC AAW71617;

XX 14-DEC-1998 (first entry)

DE Bacillus subtilis Class II EPSPS.

XX 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
 KM glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
 KM resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
 KM N-phosphonomethylglycine.

OS Bacillus subtilis.

PN US5804425-A.

PD 08-SEP-1998.

PF 07-APR-1997; 97US-0833485.

PR 13-SEP-1994; 94US-0306063.

PR 31-AUG-1990; 90US-0576537.

PR 28-AUG-1991; 91US-0749611.

PR 07-APR-1997; 97US-0833485.

XX (MONS) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

XX WPI; 1998-505657/43.

XX N-PSDB; AAV58018.

XX Glyphosate resistant 5-enolpyruvylshikimate-3-phosphate synthase -
 PT useful for characterisation of the enzyme to determine inhibition
 PT data values

XX Disclosure; Fig 18; 152pp; English.

CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
 CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
 CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
 CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
 CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
 CC produced by recombinant methods, can be used in kinetic studies to
 CC determine Ki and Km values of the enzyme for its characterisation. The
 CC enzyme is normally used for the production of 5-enolpyruvyl-3-
 CC phosphoshikimic acid in plants, and most forms of the enzyme are
 CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
 CC Inhibition data enables more accurate values of concentrations of
 CC herbicide to be used when growing the plant without being detrimental to
 CC it. This enables the plant to be grown in the presence of the herbicide,
 CC being used to inhibit the growth of undesired plants. The present
 CC sequence represents a Class II EPSPS from bacterial isolate
 CC Bacillus subtilis.

XX Sequence 428 AA;

Query Match 90.3%; Score 28; DB 19; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 11:111
 DB 241 GAVVFN 246

RESULT 12

AAE05070
 ID AAE05070 standard; Protein; 428 AA.

AC AAE05070;

DT 10-SEP-2001 (first entry)

DE Bacillus subtilis class II EPSs protein.

KM 5-enolpyruvylshikimate-3-phosphate synthase; EPSs; glyphosate;
 KM transformed bacteria; transgenic plant; herbicide.

OS Bacillus subtilis.

PN US624876-B1.

PD 19-JUN-2001.

PF 20-AUG-1998; 98US-0137440.

PR 13-SEP-1994; 94US-0306063.

PR 07-APR-1997; 97US-0833485.

PR 31-AUG-1990; 90US-0576537.

PR 28-AUG-1991; 91US-0749611.

XX (MONS) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

XX WPI: 2001-407326/43.

XX DR N-PSDB; AAD09775.

XX DNA probe capable of use in a polymerase chain reaction for identifying

XX the presence of a target genomic DNA encoding a

XX 5-enolpyruvylshikimate-3-phosphate synthase (EPSs) enzyme -

XX PS Disclosure; Fig 18; 152pp; English.

XX The present invention relates to a DNA probe capable of use in a

XX polymerase chain reaction for identifying the presence of a target

XX genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSs)

XX enzyme referred to as class II EPSs enzyme which is tolerant to

XX glyphosate. EPSs genes are useful in producing transformed bacteria

XX and transgenic plants which are tolerant to glyphosate herbicide. The

XX probe is useful for identifying the presence of a target genomic DNA

XX encoding a EPSs enzyme. The present sequence is Bacillus subtilis

XX class II EPSs protein.

XX Sequence 428 AA;

Query Match 90.3%; Score 28; DB 22; Length 428;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 11:111
 DB 241 GAVVFN 246

RESULT 13

ABB97384
 ID ABB97384 standard; Protein; 469 AA.

AC ABB97384;
 XX 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 652.

KM Human; anti-naemic; vulnerary; anti-inflammatory; immunomodulator;

KM antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI: 2002-292408/33.

XX DR N-PSDB; ABB97384.

XX An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 20; SEQ ID NO 652; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

XX novel human proteins. These were isolated from expressed sequences tags

XX (ESTs). They can be used to stimulate cell growth, to regulate

XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

XX e.g. in burn treatment, to regulate the immune system e.g. to treat

XX multiple sclerosis, to regulate activin or inhibin e.g. to treat

XX infertility, to regulate haemostasis or thrombolysis e.g. to treat

XX stroke and cancer, to screen for drugs, to treat inflammatory conditions

XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

XX Parkinson's disease. The present sequence is a protein of the invention.

XX Sequence 469 AA;

Query Match 90.3%; Score 28; DB 23; Length 469;
 Best Local Similarity 83.3%; Pred. No. 5.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVFN 6
 11:111
 DB 389 GAVVFN 394

RESULT 14

AAU16361
 ID AAU16361 standard; Protein; 497 AA.

XX Human novel secreted protein, Seq ID 1314.

XX Human; immunosuppressive; antiarthritic; antirheumatic;

XX cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;

XX neuroprotective; antibacterial; virucide; fungicide; ophthalmological;

XX vulnerary; secreted protein; rheumatoid arthritis;

XX hyperproliferative disorder; cardiovascular disorder; cardiac arrest;

XX cerebrovascular disorder; cerebral ischaemia; angiogenesis;

XX nervous system disorder; Alzheimer's disease; infection; ocular disorder;

KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.
 OS Homo sapiens.
 XX
 XX WO200155322-A2.
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01341.
 PF
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229511.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
DR WPI: 2001-488783/53.
DR N-PSDB: AAS26348.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 11: SEQ ID NO 1314; 980pp: English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capbilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention.

Query Match 90.3%; Score 28; DB 22; Length 497;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVPP 6
1-11111
Db 417 GSVVPP 422

RESULT 15
ABU55430
ID ABU55430 standard; Protein; 497 AA.
XX
AC ABU55430;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #517.
XX
XX Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
OS
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX

PF 17-JAN-2001; 2001US-0764864.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225257P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI: 2003-147444/14.
DR N-PSDB: ABX73689.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for
PT treating, inhibiting or preventing e.g. neural, immune system,
PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
PT cardiovascular or renal disorders -
XX
XX
PS Claim 11: SEQ ID NO 1314; 402pp: English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid

CC arthritis and multiple sclerosis), muscular disorders, respiratory
CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
CC cardiovascular disorders (e.g. congenital heart defects, Ebstein's
CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
CC kidney failure and end-stage renal disease), hyperproliferative disorders
CC (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g.
CC septic shock, bursitis and appendicitis), allergic reactions and
CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
CC atherosclerosis and myocardial infarction) and cancerous diseases.
CC Sequences ABU54914-ABU55699 and ABU55748 represent human novel
CC polypeptides of the invention.
XX

SO Sequence 497 AA;

Query Match 90.3%; Score 28; DB 24; Length 497;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPV 6

Db 417 GSVVPPN 422

Search completed: September 4, 2003, 21:00:31
Job time : 30.8824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 9.35294 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893d-59
Perfect score: 31
Sequence: 1 GAVVPN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	468	2 T22397	hypothetical prote
2	31	100.0	1360	2 T12064	DNA binding protei
3	30	96.8	72	2 T12130	potassium channel
4	30	96.8	72	2 T03391	potassium channel
5	30	96.8	99	1 W5BP57	gene 5.5 protein -
6	30	96.8	340	2 F95266	probable lact-fam1
7	30	96.8	587	1 W2BEC9	gene 36 protein -
8	30	96.8	587	2 T42580	gene 36 protein -
9	30	96.8	803	1 WMV994	early 94K protein
10	30	96.8	803	2 G72866	hypothetical prote
11	30	96.8	807	2 T12177	potassium channel
12	30	96.8	1891	2 T43262	calcium channel al
13	28	90.3	190	2 E84355	DNA-directed RNA p
14	28	90.3	339	2 A13609	daunorubicin resis
15	28	90.3	425	2 A87631	FMN oxidoreductase
16	28	90.3	425	2 T21835	hypothetical prote
17	28	90.3	428	2 C26532	5-enolpyruvylshik
18	28	90.3	468	2 C95405	probable sulfite o
19	28	90.3	469	2 T46929	hypothetical prote
20	28	90.3	469	2 T46930	hypothetical prote
21	28	90.3	530	2 S52215	hypothetical prote
22	28	90.3	547	2 T06758	probable galactose
23	28	90.3	569	1 A24984	ribulokinase (EC 2
24	28	90.3	569	2 A00515	L-ribulokinase (im
25	28	90.3	586	2 T08293	hypothetical prote
26	28	90.3	636	2 D84823	hypothetical prote
27	27	87.1	66	2 S03595	chr protein - phag
28	27	87.1	108	2 B43936	ORF 3' of ctsA - B
29	27	87.1	131	1 H52PA3	histone H2A.2 - fl

30	27	87.1	132	1 H52PA2	histone H2A.1 - fl
31	27	87.1	213	2 D71669	adenylate kinase (
32	27	87.1	239	2 F81704	serine esterase, p
33	27	87.1	239	2 A71552	probable lysophosp
34	27	87.1	339	2 F86683	prophage p11 prote
35	27	87.1	451	2 T32777	hypothetical prote
36	27	87.1	483	2 F64760	membrane protein p
37	27	87.1	483	2 C90677	hypothetical prote
38	27	87.1	483	2 F85527	hypothetical prote
39	27	87.1	483	2 A10547	PRP protein (impo
40	27	87.1	537	2 H90608	conserved hypochet
41	27	87.1	620	2 T50232	actin-like protein
42	27	87.1	682	2 T12968	hypothetical prote
43	27	87.1	711	2 T30107	hypothetical prote
44	27	87.1	744	2 T45943	hypothetical prote
45	27	87.1	767	1 S50594	5-methyltetrahydro

ALIGNMENTS

RESULT 1
T22397
hypothetical protein F49A5.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T22397
R:Morlmore, B.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z19560
A:Accession: T22397
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1468 <MIL>
A:Cross-references: EMBL:281542; PIDN:CAB04419.1; GSPDB:GN00023; CESP:F49A5.7
A:Experimental source: clone F49A5
C:Genetics:
A:Gene: CESP:F49A5.7
A:Map position: 5
A:Introns: 19/1; 66/1; 151/1; 174/1; 189/1; 224/3; 248/3; 297/1; 340/1

Query Match 100.0%; Score 31; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
DB 401 GAVVPN 406

RESULT 2
T12064
DNA binding protein G11-1 - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Nov-2000
C/Accession: T12064
R:Ruiz i Altaba, A.; Lee, J.D.
Submitted to the EMBL Data Library, May 1996
A:Reference number: Z17399
A:Accession: T12064
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1360 <RU1>
A:Cross-references: EMBL:U57454; NID:g3282202; PID:g3282203
C:Superfamily: g11 transforming protein
C:Keywords: zinc finger

Query Match 100.0%; Score 31; DB 2; Length 1360;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAVVPN 6
|||||

Db 1332 GAVPN 1337

RESULT 3

T12130

C:Species: Vicia faba (fava bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Dec-2002

C:Accession: T12130

A:Reference number: Z17429

A:Accession: T12130

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-72 <PHI>

A:Cross-references: EMBL:Y09749; NID:e1012614; PID:e293987

A:Experimental source: leaf

C:Superfamily: Arabidopsis potassium channel protein AKT1; ankyrin repeat homology; CAM

C:Keywords: potassium channel; potassium transport; transmembrane protein; transport pro

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 72;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 4

T03391

C:Species: Zea mays (maize)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000,

C:Accession: T03391

R:Hotl, S.; Dreyer, I.; Dietrich, P.; Becker, D.; Mueller-Roeber, B.; Hedrich, R.

Proc. Natl. Acad. Sci. U.S.A. 94, 4806-4810, 1997

A:Title: Molecular basis of plant-specific acid activation of K⁺ uptake channels.

A:Reference number: Z14922; MUID:97272307; PMID:9114073

A:Accession: T03391

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-72 <HOT>

A:Cross-references: EMBL:Y09747; NID:g2104905; PIDN:CAW0894.1; PID:g2104906

A:Experimental source: strain 155111B; coleoptile

C:Keywords: potassium channel; potassium transport; transmembrane protein

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 72;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

C:Species: phage T7

C:Date: 13-Jun-1993 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000

C:Accession: A04412; S42313

R:Dunn, J.J.; Thompson, K.

submitted to the Nucleic Acid Sequence Database, September 1982

A:Reference number: A94615

A:Accession: A04412

A:Molecule type: DNA

A:Residues: 1-99 <DUN>

R:Dunn, J.J.; Studier, F.W.

J. Mol. Biol. 166, 477-535, 1983

A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge

A:Accession: S42313

A:Molecule type: DNA

A:Residues: 1-99 <DUN>

A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAW4414.1; PID:g15593

C:Genetics: the authors did not translate the codon for residue 1

A:Gene: 5.5

A:Map position: 42.20-43.46

C:Superfamily: phage T7 gene 5.5 protein

Query Match

Best Local Similarity 96.8%; Score 30; DB 1; Length 99;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 6

F95266

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: F95266

R:Barneet, M.U.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95262; MUID:2136509; PMID:11481432

A:Accession: F95266

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <KUR>

A:Cross-references: GB:AE006469; PIDN:AK64696.1; PID:g14523096; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pela, D.; Chain, P.; Coyle, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Reference number: A96039; MUID:21368234; PMID:11474104

C:Contents: annotation

C:Genetics:

A:Gene: Sma0078

C:Superfamily: lac repressor

Query Match

Best Local Similarity 96.8%; Score 30; DB 2; Length 340;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

C:Species: equine herpesvirus 1

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: B36799

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to Genbank, March 1992

A:Reference number: A36805

A:Accession: B36799

A:Molecule type: DNA

A:Residues: 1-587 <TEL>

A:Cross-references: GB:M86664; NID:g330791; PIDN:AA802472.1; PID:g330829
R:Teiford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 36
C:Superfamily: varicella-zoster virus gene 34 protein

Query Match 96.8%; Score 30; DB 1; Length 587;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
11:111
Db 413 GATVPN 418

RESULT 8
T42580
gene 36 protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42580
R:Teiford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.

A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: Z22173; MUID:98264497; PMID:9603335
A:Accession: T42580
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-587 <REL>

A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:MAC59553.1; PID:g2605961
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 36
C:Superfamily: varicella-zoster virus gene 34 protein

Query Match 96.8%; Score 30; DB 2; Length 587;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
11:111
Db 413 GATVPN 418

RESULT 9

WMNV94

early 94K protein - Autographa californica nuclear polyhedrosis virus (strain L-1)

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: B27840
R:Friesen, P.D.; Miller, L.K.

J. Virol. 61, 2264-2272, 1987
A:Title: Divergent transcription of early 35- and 94-kilodalton protein genes encoded by

A:Reference number: A93026; MUID:87226411; PMID:3035225
A:Accession: B27840
A:Molecule type: DNA
A:Residues: 1-803 <PRI>

A:Cross-references: GB:M16821; NID:g332445; PIDN:AAA6702.1; PID:g332446
C:Superfamily: Autographa californica nuclear polyhedrosis virus early 94K protein
C:Keywords: early protein

Query Match 96.8%; Score 30; DB 1; Length 803;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
11:111

Db 412 GATVPN 417

RESULT 10

C72866

hypothetical protein - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-Jan-2000
C:Accession: G72866
R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Reber, M.; Possee, R.D.

Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: G72866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <NR>

A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66764.1; PID:g559203
C:Genetics:
A:Gene: AC-94K
C:Superfamily: Autographa californica nuclear polyhedrosis virus early 94K protein

Query Match 96.8%; Score 30; DB 2; Length 803;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
11:111
Db 412 GATVPN 417

RESULT 11

T12177

potassium channel protein - fava bean (fragment)

C:Species: Vicia faba (fava bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Dec-2002
C:Accession: T12177
R:Ache, P.; Wohlfarth, T.; Hedrich, R.; Becker, D.

submitted to the EMBL Data Library, January 1997
A:Reference number: Z17443
A:Accession: T12177
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-807 <ACH>

A:Cross-references: EMBL:Y10579; NID:e1015303; PID:e321441
A:Experimental source: cotyledon; clone VKC1
C:Superfamily: Arabidopsis potassium channel protein AKT1; ankyrin repeat homology; C
C:Keywords: potassium channel; transmembrane protein; transport protein; voltage-gate
F;550-582/Domain: ankyrin repeat homology <ANR>

Query Match 96.8%; Score 30; DB 2; Length 807;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
11:111
Db 230 GATVPN 235

RESULT 12

T43262

calcium channel alpha-1 chain, L-type - Stylophora pistillata

C:Species: Stylophora pistillata
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T43262
R:Zoccola, D.; Tamburte, E.; Senegas-Balas, F.; Michiels, J.F.; Falla, J.P.; Jaubert

Gene 227, 157-167, 1999
A:Title: Cloning of a calcium channel alpha subunit from the reef-building coral, St
A:Reference number: Z22375; MUID:99148007; PMID:10023047
A:Accession: T43262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A;Residues: 1-1891 <ZOC>
A;Cross-references: EMBL:U64465; NID:q4204977; PID:q4204978; PIDN:AA011470.1
C;Genetics:
A;Gene: CACHL
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 96.88; Score 30; DB 2; Length 1891;
Best Local Similarity 83.38; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
Db 744 GAVVFN 749

RESULT 13

E84355
DNA-directed RNA polymerase subunit E' [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: E84355

R;N: W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic,
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: E84355

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-190 <STO>

A;Cross-references: GB:AE004437; NID:q10581476; PIDN:AAG20209.1; GSPDB:GN00138

C;Genetics:

A;Gene: rpoE'

C;Superfamily: DNA-directed RNA polymerase subunit E

Query Match 90.38; Score 28; DB 2; Length 190;
Best Local Similarity 83.38; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
Db 57 GAVVFN 62

RESULT 14

A13609

daunorubicin resistance ATP-binding protein drrA BME110802 [imported] - Brucella melitensis

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C;Accession: A13609

R;DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessie

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: A13609

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-339 <KUR>

A;Cross-references: GB:AE008918; PIDN:AA154044.1; PID:q17984998; GSPDB:GN00191

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME110802

A;Map position: 11

C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 90.38; Score 28; DB 2; Length 339;
Best Local Similarity 83.38; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6

Db 31 GAVVFN 36

RESULT 15

A87631

FMN oxidoreductase CC3083 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: A87631

R;Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87631

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-425 <STO>

A;Cross-references: GB:AE005673; NID:q13424735; PIDN:AAK25045.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC3083

Query Match 90.38; Score 28; DB 2; Length 425;
Best Local Similarity 83.38; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
Db 18 GAVVFN 23

Search completed: September 4, 2003, 21:10:48
Job time : 9.35294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 ; Search time 5.20588 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-59
Perfect score: 31
Sequence: 1 GAVVFN 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No: is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	1 GLI1_XENLA	Q91690 xenopus lae
2	30	96.8	1 V55_BPT7	P03787 bacterioph
3	30	96.8	1 UL25_HSVB	P28928 equine herp
4	30	96.8	1 VE94_NPVAC	P08161 autographa
5	28	90.3	1 PEB1_DROME	Q9u615 drosophila
6	28	90.3	1 AROA_BACSU	P20691 bacillus su
7	28	90.3	1 AROA_SALTI	P58542 salmonella
8	28	90.3	1 AROA_SALTI	P58542 salmonella
9	27	87.1	1 VDRH_BP186	P21680 bacterioph
10	27	87.1	1 H2A2_SCHPO	P04910 schizosacch
11	27	87.1	1 H2A2_SCHPO	P04910 schizosacch
12	27	87.1	1 ML1A_XENLA	P51048 xenopus lae
13	27	87.1	1 KAD_RICPR	Q9ZCS6 rickettsia
14	27	87.1	1 PRPD_ECO57	Q8X693 escherichia
15	27	87.1	1 PRPD_ECOLI	P77243 escherichia
16	27	87.1	1 PRPD_SALTI	O8Z903 salmonella
17	27	87.1	1 PRPD_SALTI	P74840 salmonella
18	27	87.1	1 ARBP_SCHPO	Q9US87 schizosacch
19	27	87.1	1 METE_YEAST	P05694 saccharomyc
20	27	87.1	1 TTXA_DROME	P42282 drosophila
21	27	87.1	1 DNL1_RAT	Q91YH8 ratus norv
22	27	87.1	1 POLG_PYVN	P18227 p genome po
23	26	83.9	1 H2A1_WHEAT	P02275 triticum ae
24	26	83.9	1 H2A2_WHEAT	P02276 triticum ae
25	26	83.9	1 H2A3_WHEAT	P02277 triticum ae
26	26	83.9	1 YWOC_BACSU	P94573 bacillus su
27	26	83.9	1 RL3_BACST	P28600 bacillus st
28	26	83.9	1 Y237_THEAC	Q9H1J1 thermoplasm
29	26	83.9	1 RM21_ARATH	P81940 arabidopsis
30	26	83.9	1 DRN1_HUMAN	P24855 homo sapien
31	26	83.9	1 M48A_MOUSE	Q99N10 mus musculu
32	26	83.9	1 E434_ADOB9	P89083 human adeno
33	26	83.9	1 ADD_TREPA	O83085 treponema p

34	26	83.9	320 1	CHIX_PEA	P36907 pisum sativ
35	26	83.9	333 1	CCPA_STRMU	O07339 streptococc
36	26	83.9	336 1	CH12_ORISA	P25765 oryza sativ
37	26	83.9	345 1	HAIF_CHICK	P15979 gallus gall
38	26	83.9	353 1	DPE1_HUMAN	Q92782 homo sapien
39	26	83.9	353 1	PSBA_PROHO	P15191 prochloroth
40	26	83.9	356 1	REOC_MOUSE	P51759 cyanothece
41	26	83.9	356 1	PSBA_PROHO	P51759 cyanothece
42	26	83.9	360 1	PSB1_ANASP	P46242 anabaena sp
43	26	83.9	360 1	PSB1_STREL	P35876 synechococc
44	26	83.9	360 1	PSB1_SYNP7	P04996 synechococc
45	26	83.9	360 1	PSB1_SYNVU	P51765 synechococc

ALIGNMENTS

```

RESULT 1
ID      GLI1_XENLA      STANDARD:      PRT: 1360 AA.
AC      091690:
DT      01-NOV-1997 (Rel. 35, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Zinc finger protein GLI1 (GLI-1) (Fragment).
GN      GLI1.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9735968; PubMed=9216996;
RA      Lee J., Platt K.A., Censullo P., Ruiz i Altaba A.;
RT      "Gli1 is a target of Sonic hedgehog that induces ventral neural tube
RT      development."
RL      Development 124:2537-2552(1997).
RN      [2]
RP      REVISIONS.
RA      Lee J., Platt K.A., Censullo P., Ruiz i Altaba A.;
RT      Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: IMPLICATED IN THE TRANSDUCTION OF SHH SIGNAL (BY
CC      SIMILARITY).
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC      PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U57454; AAC24946.1;
DR      PIR: T12064; T12064.
DR      HSSP: P08151; 2GLI.
DR      InterPro: IPR007087; znf_C2H2.
DR      Pfam: PF00096; zf_C2H2_5.
DR      SMART: SM00355; znf_C2H2_5.
DR      PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR      PROSITE: PS50157; ZINC_FINGER_C2H2_2; 5.
KW      Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;
KW      Nuclear protein; Repeat.
FT      ZN_FING      250      275      C2H2-TYPE.
FT      ZN_FING      316      340      C2H2-TYPE.
FT      ZN_FING      346      371      C2H2-TYPE.
FT      ZN_FING      377      402      C2H2-TYPE.
FT      NON_TER      1360      1360
SQ      SEQUENCE      1360 AA; 149422 MW; 5A32B806794EC2D CRC64;

```

Query Match 100.0%; Score 31; DB 1; Length 1360.
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
 DB 1332 GAVPN 1337

RESULT 2
 V55_BPT7
 ID V55_BPT7 STANDARD; PRT; 169 AA.
 AC P03787;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gene 5.5 protein and probable fusion gene 5.5-5.7 protein.
 GN 5.5.
 OS Bacteriophage T7.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like viruses.
 OX NCBI_Taxid=10760;
 RN [1]
 RX MEDLINE=63241725; PubMed=6864790;
 RA Dunn J.J., Studier F.W.;
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
 RT locations of T7 genetic elements.";
 RL J. Mol. Biol. 166:477-535(1983).
 RN [2]
 RP REVISIONS.
 RA Dunn J.J., Studier F.W.;
 RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
 CC -1- MISCELLANEOUS: AN AMBER MUTATION IN GENE 5.5 RESULTS IN THE LOSS
 CC OF TWO PROTEINS. BY ANALOGY WITH THE GENE 10B PRODUCT, IT IS
 CC BELIEVED THAT THE SECOND PROTEIN IS PRODUCED BY A -1 FRAME-SHIFT
 CC DURING THE TRANSLATION OF GENE 5.5. THIS FRAME-SHIFT MAY OCCUR
 CC ANYWHERE WITHIN THE REGION CORRESPONDING TO RESIDUES 85-99 OF THE
 CC SEQUENCE SHOWN. THUS A PORTION OF THESE RESIDUES MAY BE REPLACED
 CC BY A PORTION OF THE SEGMENT GLPQTISCSNGVYL IN THE ACTUAL GENE
 CC 5.5-5.7 PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; V01146; CAA24415.1; -
 DR EMBL; V01146; CAA24414.1; ALT-TERM.
 DR EMBL; V01146; CAA24416.1; ALT-INIT.
 FT CHAIN 1 99 GENE PROTEIN 5.5.
 FT CHAIN 1 169 FUSION GENE PROTEIN 5.5-5.7.
 SQ SEQUENCE 169 AA; 18713 MW; 4AF673FDD122E8A99 CRC64;

Query Match 96.8%; Score 30; DB 1; Length 169;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
 DB 39 GAVPN 44

RESULT 3
 UL25_HSVB
 ID UL25_HSVB STANDARD; PRT; 587 AA.
 AC P28928;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)

DE Viron protein UL25.
 GN UL25 OR 36.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_Taxid=31520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 CC -1- FUNCTION: VIRION PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL25,
 CC EHV-1 36, EBV BFRF, HCMV UL77, ILTV ORF2, AND VZV 34.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; M86664; AAB02472.1; -
 DR PIR; B36799; WZBEC9.
 DR InterPro; IPR002493; UL25.
 DR Pfam; PF01499; UL25; 1.
 SQ SEQUENCE 587 AA; 63692 MW; 6BE2F5FAB34CA3C8 CRC64;

Query Match 96.8%; Score 30; DB 1; Length 587;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
 DB 413 GAVPN 418

RESULT 4
 VE94_NPVAC
 ID VE94_NPVAC STANDARD; PRT; 803 AA.
 AC P08161;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Early 94 kDa protein.
 OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=LI;
 RC MEDLINE=87226411; PubMed=3035225;
 RX Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "Divergent transcription of early 35- and 94-kilodalton protein genes
 RT encoded by the HindIII K genome fragment of the baculovirus
 RT Autographa californica nuclear polyhedrosis virus.";
 RL J. Virol. 61:2264-2272(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=C6;
 RC MEDLINE=94303173; PubMed=8030224;
 RX Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 RT polyhedrosis virus.";
 RL Virology 202:586-605(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M16821: AAA46702.1; -
 DR EMBL: L22858: AAA66764.1; -
 DR PIR: B27840: MNV94.
 DR PIR: G72866: G72866.
 KW Early protein.
 FT CONFLICT 457 E -> K (IN REF. 1).
 SQ SEQUENCE 803 AA: 94540 MM: A89964321382097D CRC64;

Query Match
 Best Local Similarity 96.8%; Score 30; DB 1; Length 803;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPP 6
 DB 412 GATVPP 417

RESULT 5
 PEBL_DROME STANDARD; PRT; 377 AA.
 ID PEBL_DROME
 AC Q9U6L5; Q9W0M1;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ejaculatory bulb specific protein I precursor (PEB-me).
 GN PEB OR BCDNA:GH06048 OR CG2668.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RX MEDLINE-20196006; PubMed-10731132;
 RC STRAIN-Berkeley;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaziel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agyarani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklav G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP REVISIONS.
 RC STRAIN-Berkeley;
 RX MEDLINE-22426069; PubMed-12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertan B.P.,
 RA Belencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley; TISSUE-Head;
 RX MEDLINE-20196012; PubMed-10731138;
 RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
 RA Stapleton M., Harvey D.A.;
 RT "A *Drosophila* complementary DNA resource.";
 RL Science 287:2222-2224(2000).
 RN [4]
 RP SEQUENCE OF 77-90 AND 101-109, FUNCTION, SUBCELLULAR LOCATION, TISSUE
 RP SPECIFICITY, AND VARIANT.
 RC STRAIN-Canton-S;
 RX MEDLINE-21167515; PubMed-11267893;
 RA Lung O., Wolfner M.F.;
 RT "Identification and characterization of the major *Drosophila*
 RT melanogaster mating plug protein.";
 RL Insect Biochem. Mol. Biol. 31:343-351(2001).
 CC -1- FUNCTION: Major protein component of the posterior mating plug.
 CC Accessory gland proteins constitute, or are required for formation
 CC of the anterior mating plug. Posterior mating plug forms before
 CC sperm transfer and the anterior mating plug is formed after the
 CC start of mating.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Specifically expressed in the ejaculatory bulb
 CC and seminal fluid. Detected in mated females 3 minutes after the
 CC start of mating, and for at least 3 hours after the start of
 CC mating.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE003466; AAF47319.2; -
 DR EMBL: AF184225; AAD5736.1; -
 DR FlyBase: FBgn0004181; Pdb.
 KW Behavior: Signal; Polymorphism.
 FT SIGNAL 1 20
 FT CHAIN 11 377
 FT DOMAIN 21 256
 FT DOMAIN 159 377
 FT VARIANT 84 84
 FT SEQUENCE 377 AA: 37819 MM: D020F027BD96B221 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 1; Length 377;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPP 6
 DB 261 GSVPV 266

RESULT 6
 AROA_BACSU STANDARD: PRT: 428 AA.
 AC P20691;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 AROA.
 GN Bacillus subtilis.
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_Taxid=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87191994; PubMed=3106153;
 RA Henner D.J., Band L., Flagg G., Chen E.;
 RT "The organization and nucleotide sequence of the Bacillus subtilis
 hist. tyra and aroE genes.";
 RL Gene 49:147-152(1986).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Broillet S., Burschl C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Ghin S.Y., Glaser P., Gottfau A., Gollightly E.J., Grandi G.,
 Guseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaerr-blanchard M., Klein C.,
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 Prescann E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 Sekiguchi J., Sekowska A., Serr S.U., Serr P., Shin B.S., Soldo B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 Takeuchi V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 sixth step.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the EPSP synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M80245; AAA20869.1; -

DR EMBL, Z99115; CAB14176.1; -
 DR PIR, C26532; C26532.
 DR Subtilist; BG10294; aroE;
 DR HAMAP, ME_00210; -, 1.
 DR InterPro, IPR006264; AroA.
 DR InterPro, IPR001986; EPSP_synthase.
 DR Pfam, PF00275; EPSP_synthase; 1.
 DR TIGRfams; TIGR01356; aroA; 1.
 DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
 DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome;
 SQ SEQUENCE 428 AA; 45240 MW; DE3F7B96E761CB40 CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 428;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAVPN 6
 DB 241 GAVPN 246
 RESULT 7
 AROA_SALTI STANDARD: PRT: 568 AA.
 AC P58542;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE L-ribulokinase (EC 2.7.1.16).
 GN AROA OR STY0120 OR T0107.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=601;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahin M.,
 Baker S., Baaham D., Brooks K., Chillingworth T., Conerton P.,
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 Feltwell T., Hamlin N., Hogue A., Hien T.T., Holtroyd S., Jaggels K.,
 Krogh A., Larsen T.S., Leach S., Moute S., O'Geora P., Parry C.,
 Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liu S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 Burland V., Kodoyanni V., Schwartz D.C., Blattner D.C.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ribulose = ADP + L-ribulose 5-
 phosphate.
 CC -1- PATHWAY: L-arabinose catabolism; second step.
 CC -1- SIMILARITY: Belongs to the ribulokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC

```

DR EMBL; AL627265; CAD01260.1; -
DR EMBL; AE016834; AAC67839.1; -
DR HAMAP; MF_00520; -; 1.
DR InterPro: IPR000577; FGGY_kin.
DR InterPro: IPR005929; L-ribulokin.
DR Pfam; PF00370; FGGY_1.
DR Pfam; PF02782; FGGY_C; 1.
DR TIGRfams; TIGR01234; L-ribulokinase; 1.
DR Transferase; Kinase; Arabinose catabolism; Complete proteome.
KW INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 568 AA; 61643 MW; A1515DB7249A0F75 CRC64;

Query Match          90.3%; Score 28; DB 1; Length 568;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
DB 321 GSAVFN 326

RESULT 8
ID ARAB_SALTY STANDARD; PRT; 568 AA.
AC P06188;
DT 01-JAN-1988 (Rel. 06, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-ribulokinase (EC 2.7.1.16).
GN ARAB OR STM0103.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85232044; Pubmed=2989100;
RA Lin H.-C., Lei S.-P., Wilcox G.;
RT "The arabid operon of Salmonella typhimurium ltr2. I. Nucleotide
RT sequence of arab and primary structure of its product,
RT ribulokinase."
RL Gene 34:111-122(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SCS1412 / ATCC 700720;
RX MEDLINE=21534948; Pubmed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mufvany E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -|- CATALYTIC ACTIVITY: ATP + L-ribulose = ADP + L-ribulose 5-
CC phosphate.
CC -|- PATHWAY: L-arabinose catabolism; second step.
CC -|- SIMILARITY: Belongs to the ribulokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11047; AAA27023.1; -
DR EMBL; AE008698; AAL19067.1; -
DR PIR; A24984; A24984.
DR SLYGene; SG10013; arab.
DR HAMAP; MF_00520; -; 1.
DR InterPro: IPR000577; FGGY_kin.

```

```

DR InterPro: IPR005929; L-ribulokin.
DR Pfam; PF00370; FGGY_1.
DR Pfam; PF02782; FGGY_C; 1.
DR TIGRfams; TIGR01234; L-ribulokinase; 1.
DR Transferase; Kinase; Arabinose catabolism; Complete proteome.
KW INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 568 AA; 61620 MW; BID838C37BCA134 CRC64;

Query Match          90.3%; Score 28; DB 1; Length 568;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
DB 321 GSAVFN 326

RESULT 9
ID VDRH_BP186 STANDARD; PRT; 66 AA.
AC P21680;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE DHR protein.
GN DHR OR CP78.
OS Bacteriophage 186.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC P2-like viruses.
OX NCBI_TaxID=29252;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8919647; Pubmed=2704042;
RA Richardson H., Puspurs A., Egan J.B.;
RT "Control of gene expression in the P2-related temperate coliphage
RT 186. VI. Sequence analysis of the early lytic region."
RL J. Mol. Biol. 206:251-255(1989).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=8919651; Pubmed=2704043;
RA Richardson H., Egan J.B.;
RT "DNA replication studies with coliphage 186. II. Depression of host
RT replication by a 186 gene."
RL J. Mol. Biol. 206:59-68(1989).
CC -|- FUNCTION: INVOLVED IN THE DEPRESSION OF HOST DNA REPLICATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15001; CAA33108.1; -
DR EMBL; U32222; AAC34180.1; -
DR PIR; S03595; S03595.
KW Early protein.
SQ SEQUENCE 66 AA; 7531 MW; 64E93005DA7C0F7 CRC64;

Query Match          87.1%; Score 27; DB 1; Length 66;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVVFN 6
DB 11 GAMIPN 16

RESULT 10
ID H2A2_SCHPO STANDARD; PRT; 130 AA.
ID H2A2_SCHPO
AC P04910;

```

DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A-beta (H2A.2).
 GN H2A2 OR SPAC19G12.06c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NCBI_TaxID=4896;
 RX MEDLINE-86135992; PubMed=4092687;
 RX MEDLINE-86135992; PubMed=4092687;
 RX "Histone gene organization of fission yeast: a common upstream
 sequence."
 RT EMBO J. 4:3531-3538(1985).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE-21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Dague R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 RT -1- SUBUNIT: The nucleosome is an octamer containing two molecules
 of each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
 bp of DNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the histone H2A family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL, M1500; AAA35310.1;
 CC EMBL, X05221; CAA28849.1;
 CC EMBL, Z97209; CAB10117.1;
 CC PIR: C27399; HSP2A3.
 DR GeneDB_spombe; SPAC19G12.06c; -

DR InterPro: IPR004822; Histone core.
 DR InterPro: IPR002119; Histone_H2A.
 DR Pfam: PF00125; histone_1.
 DR PRINTS: PR00620; HISTONEH2A.
 DR ProDom: PD000522; Histone_H2A.1.
 DR SMART: SM00414; H2A.1.
 DR PROSITE: PS00046; HISTONE_H2A.1.
 KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
 KW Multigene family.
 FT INIT MET
 SQ SEQUENCE 130 AA; 13645 MW; 1BD5C86BC8692A04 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 130;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 GY 1 GAYVFN 6
 DB 106 GGAVFN 111
 RESULT 11
 H2A1_SCHPO STANDARD; PRT; 131 AA.
 ID H2A1_SCHPO
 AC P04909;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histone H2A-alpha (H2A.1).
 GN H2A1 OR SPC622.08C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NCBI_TaxID=4896;
 RX MEDLINE-86135992; PubMed=4092687;
 RX MEDLINE-86135992; PubMed=4092687;
 RX "Histone gene organization of fission yeast: a common upstream
 sequence."
 RT EMBO J. 4:3531-3538(1985).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE-21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of *Schizosaccharomyces pombe*."
 RL Nature 415:871-880(2002).
 CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules
 CC each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
 CC bp of DNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the histone H2A family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M1494; AAA35311.1; -
 DR EMBL: X05220; CAA28848.1; -
 DR EMBL: AF033127; CAA21864.1; -
 DR PIR: B27399; HSZPR2.
 DR GeneDB: SPombe; SPC622.08c; -
 DR InterPro: IPR004822; Histone_core.
 DR InterPro: IPR002119; Histone_H2A.
 DR Pfam: PF00125; histone.1.
 DR PRINTS: PR00620; HISTONEH2A.
 DR PRODOM: PD000522; Histone_H2A; 1.
 DR SMART: SM00414; H2A; 1.
 DR PROSITE: PS00046; HISTONE_H2A; 1.
 KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
 KW Multigene family.
 FT INIT_MER 0
 FT CONFLICT 123 123 R -> G (IN REF. 1).
 SQ SEQUENCE 131 AA; 13747 MW; D1AFA80F580C3273 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 131;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAVVPP 6
 Db 106 GGAVPP 111
 RESULT 12
 MLIA_XENLA STANDARD; PRT; 153 AA.
 ID MLIA_XENLA
 AC P51048;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Melatonin receptor type 1A X2.0 (MEL-1A-R) (Fragment).
 OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96073557; PubMed-7576645;
 RA Reppert S.M., Weaver D.R., Cassone V.M., Godson C.,
 RA Kolakowski L.F. Jr.;
 RT "Melatonin receptors are for the birds: molecular analysis of two
 RT receptor subtypes differentially expressed in chick brain.";
 RL Neuron 15:1003-1015(1995).
 CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR MELATONIN. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY PERTUSSIS TOXIN SENSITIVE G PROTEINS
 CC THAT INHIBITS ADENYLATE CYCLASE ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U31826; AAA92500.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODOPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; PARTIAL.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT NON_TER 1
 FT DOMAIN 1 1
 FT TRANSMEM 13 33 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 34 57 4 (POTENTIAL).
 FT TRANSMEM 58 78 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 79 112 5 (POTENTIAL).
 FT DOMAIN 113 133 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 134 145 6 (POTENTIAL).
 FT TRANSMEM 146 >153 EXTRACELLULAR (POTENTIAL).
 FT NON_TER 153 153 7 (POTENTIAL).
 SQ SEQUENCE 153 AA; 17665 MW; ED0421422B20F27 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 153;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAVVPP 6
 Db 27 GATLPP 32
 RESULT 13
 KAD_RICPR STANDARD; PRT; 213 AA.
 ID KAD_RICPR
 AC Q9ZCS6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase).
 GN ADK OR RP638.
 OS *Rickettsia prowazekii*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Madrid E;
 RX MEDLINE-99039499; PubMed-9823893;
 RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
 RA Sichterlik-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of *Rickettsia prowazekii* and the origin of
 RT mitochondria.";
 RL Nature 396:133-140(1998).
 CC -1- FUNCTION: THIS SMALL, UBIQUITOUS ENZYME IS ESSENTIAL FOR
 CC MAINTENANCE AND CELL GROWTH.
 CC -1- CATALYTIC ACTIVITY: ATP + AMP -> ADP + ADP.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the adenylate kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: AJ235272; CA15078.1;
 DR PIR: D71669; D71669.
 DR HSSP: P07170; IAKY.
 DR HAMAP: MF_00235; 1;
 DR InterPro: IPR006259; Adenyl_kin_sub.
 DR InterPro: IPR00850; Adenylate_kin.
 DR Pfam: PF00406; Adk_1.
 DR Pfam: PF05191; Adk_1ld; 1.
 DR PRINTS: PR00094; ADENYLKINASE.
 DR PRODOM: PD000657; Adenylate_kin; 1.
 DR TIGRfam: TIGR01351; adk; 1.
 DR PROSITE: PS00113; ADENYLATE_KINASE; 1.
 DR Transferrase: Kinase; ATP-binding; Complete proteome.
 KW NP_BIND 7
 FT ATP (BY SIMILARITY).
 SQ SEQUENCE 213 AA; 24503 MW; 23CBD806A1BEEC19 CRC64;

Query Match
 Best Local Similarity 87.1%; Score 27; DB 1; Length 213;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVPN 6
 Db 56 GALPN 61

RESULT 14
 PRPD_ECO57
 ID PRPD_ECO57 STANDARD; PRT; 482 AA.
 AC Q8X693;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 2-methylcitrate dehydratase (EC 4.2.1.79).
 GN PRPD OR 20429 OR EC50387.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocitis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 RA Hida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Catalyzes the dehydration of 2-methylcitrate to 2-
 methyl-cis-aconitate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-hydroxybutane-1,2,3-tricarboxylate + 2-
 but-2-ene-1,2,3-tricarboxylate + H(2)O.
 CC -1- COFACTOR: Contains one 2Fe-2S cluster (By similarity).
 CC -1- PATHWAY: Propionate catabolism; 2-methylcitric acid cycle; third
 step.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MMGE / PRPD FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).

DR EMBL: AF005212; AAC54692.1;
 DR EMBL: AP002551; BAB33810.1;
 DR PIR: G90677; C90677.
 DR PIR: F85527; F85527.
 DR InterPro: IPR005656; MMGE_PRPD.
 DR Pfam: PF03972; MMGE_PRPD; 1.
 KW Lyase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S; Complete proteome.
 FT INT_MET 0
 FT BY SIMILARITY
 SQ SEQUENCE 482 AA; 53820 MW; 19714CD931C08227 CRC64;

Query Match
 Best Local Similarity 87.1%; Score 27; DB 1; Length 482;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAVPN 6
 Db 66 GTVVPN 71

RESULT 15
 PRPD_ECOLI
 ID PRPD_ECOLI STANDARD; PRT; 482 AA.
 AC P77243;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 2-methylcitrate dehydratase (EC 4.2.1.79).
 GN PRPD OR B0334.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
 RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
 RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 [3]
 RP CHARACTERIZATION, AND SEQUENCE OF 1-12.
 RX MEDLINE=21642584; PubMed=11782506;
 RA Blank L., Green J., Guest J.R.;
 RT "A clone of Escherichia coli is a 2-methylcitrate dehydratase (Prpd) that
 can use citrate and isocitrate as substrates";
 RL Microbiology 148:133-146(2002).
 CC -1- FUNCTION: Catalyzes the dehydration of 2-methylcitrate to 2-
 methyl-cis-aconitate. Also seems to be responsible for the
 CC residual aconitase activity of the acnAB-null strain.
 CC -1- CATALYTIC ACTIVITY: 2-hydroxybutane-1,2,3-tricarboxylate + (2)-
 but-2-ene-1,2,3-tricarboxylate + H(2)O.
 CC -1- COFACTOR: Contains one 2Fe-2S cluster.
 CC -1- PATHWAY: Propionate catabolism; 2-methylcitric acid cycle; third
 step.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: BELONGS TO THE MMGE / PRPD FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE000140; AAC73437.1; -
DR EMBL: U73857; AAB18058.1; -
DR PIR: F64760; F64760.
DR Ecocore: EG13603; PRPD.
DR InterPro: IPR005656; MmgE_PRPD.
DR Pfam: PF03972; MmgE_PRPD: 1.
KW Lyase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S; Complete proteome.
FT INIT_MET 0
SQ SEQUENCE 482 AA; 53820 MW; A7AA5CD391B82C5 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 482;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAVVPN 6
I I I I I
Db 66 GTVVPN 71

Search completed: September 4, 2003, 21:01:43
Job time : 5.20588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 24.6176 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-59

Perfect score: 31

Sequence: 1 GAVVPPN 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPTRMBL_23:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	31	100.0	174	2	092G06	092G06 staphylococ
2	31	100.0	224	10	08S708	08S708 oryza sativ
3	31	100.0	468	5	045534	045534 caenorhabdi
4	30	96.8	72	10	004236	004236 vicia faba
5	30	96.8	72	10	004241	004241 zea mays (m
6	30	96.8	108	13	042199	042199 ictalurus p
7	30	96.8	340	16	0931A6	0931A6 rhizobium m
8	30	96.8	411	16	08XTH4	08XTH4 raietonia s
9	30	96.8	587	12	039279	039279 equine hept
10	30	96.8	803	12	024538	024538 rachipiusia
11	30	96.8	807	10	024538	024538 vicia faba
12	30	96.8	830	10	09FY04	09FY04 populus tre
13	30	96.8	832	10	092PI4	092PI4 samanea sam
14	30	96.8	849	10	09SM12	09SM12 zea mays (m
15	30	96.8	1891	5	097017	097017 stylophora
16	29	93.5	867	2	09KIT1	09KIT1 porphyromon

17	28	90.3	90	13	09YH63	09YH63 gadus morhu
18	28	90.3	106	6	019066	019066 sus scrofa
19	28	90.3	116	13	09Y613	09Y613 gadus morhu
20	28	90.3	171	12	091F49	091F49 cyndidium m
21	28	90.3	190	17	09HML1	09HML1 halobacteri
22	28	90.3	221	16	09L1Y2	09L1Y2 streptomyce
23	28	90.3	232	16	08XXH3	08XXH3 raietonia s
24	28	90.3	339	16	08YBT4	08YBT4 bruceella me
25	28	90.3	377	5	09U615	09U615 drosophila
26	28	90.3	383	16	0982E2	0982E2 rhizobium l
27	28	90.3	421	10	09EX57	09EX57 rhizobium l
28	28	90.3	425	16	09A3W8	09A3W8 caulobacter
29	28	90.3	443	2	092422	092422 pseudomonas
30	28	90.3	448	16	08ETM6	08ETM6 shewanella
31	28	90.3	456	12	08UTB8	08UTB8 human papil
32	28	90.3	468	16	092XU3	092XU3 rhizobium m
33	28	90.3	469	4	09NS53	09NS53 homo sapien
34	28	90.3	469	4	09NS52	09NS52 homo sapien
35	28	90.3	483	16	08XRJ5	08XRJ5 raietonia s
36	28	90.3	485	10	08S0T3	08S0T3 oryza sativ
37	28	90.3	489	16	08XJ76	08XJ76 clostridium
38	28	90.3	528	5	P90862	P90862 caenorhabdi
39	28	90.3	534	12	083419	083419 pseudorabdi
40	28	90.3	540	12	085026	085026 pseudorabdi
41	28	90.3	547	10	09SVX6	09SVX6 arabidopsis
42	28	90.3	565	5	0814V6	0814V6 caenorhabdi
43	28	90.3	586	17	051980	051980 halobacteri
44	28	90.3	588	17	08P2R6	08P2R6 methanosarc
45	28	90.3	636	10	004198	004198 arabidopsis

ALIGNMENTS

RESULT 1

092G06 PRELIMINARY; PRT: 174 AA.
ID 092G06
AC 092G06;
DT 01-MAY-1999 (TREMUREL 10, Created)
DT 01-MAY-1999 (TREMUREL 10, Last sequence update)
DT 01-MAR-2003 (TREMUREL 23, Last annotation update)
DE TYPE-1 signal peptidase StpA.
GN StpA.
OS Staphylococcus carnosus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1281;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW300;
RA Matzen A., Frendl R.;
RT "Cloning and characterization of the type-1 signal peptidase of
RT Staphylococcus carnosus."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF089862; AAD09010.1; -;
DR HSSP; P00803; 1B12.
DR InterPro; IPR000508; SigPase.
DR Pfam; PF00461; Peptidase_S26; 1.
SQ SEQUENCE 174 AA; 19931 MW; 965B7A8EDB818C2B CRC64;

Query Match 100.0%; Score 31; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPPN 6
Db 28 GAVVPPN 33

RESULT 2

ID 08S708 PRELIMINARY; PRT: 234 AA.
AC 08S708;
DT 01-JUN-2002 (TREMUREL 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 GN OSJNBA0091P11.16.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaceae; Oryza.
 NCBI_TaxID=39947;
 RX STRAIN=cv. Nipponbare;
 RA Buesl C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Gansberger K., Benner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tsirlin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
 RA Vanaken S.E., Uterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBA0091P11 genomic sequence."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC073556; AAL84306.1; -.
 DR Gremene; Q85708; -.
 DR InterPro: IPR006662; ThioRed.
 DR InterPro: IPR006663; ThioRedox_dom2.
 DR Pfam: PF00085; ThioRed; 1.
 SQ SEQUENCE 234 AA; 26253 MW; B6FCC2C8293E811C CRC64;
 QY 1 GAVVPN 6
 Db 227 GAVVPN 232
 Query Match 100.0%; Score 31; DB 10; Length 234;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
 ID 045534 PRELIMINARY; PRT; 468 AA.
 AC 045534;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 DE F49A5.7 protein.
 GN F49A5.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RX SEQUENCE FROM N.A.
 RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81542; CAB04419.1; -.
 DR HSSP; P05451; 10DD.
 DR WormPep; F49A5.7; CE16070.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 2.
 DR SMART; SM00034; CLECT; 3.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_2; 2.
 SQ SEQUENCE 468 AA; 52464 MW; 71E731E29EF07118 CRC64;
 Query Match 100.0%; Score 31; DB 5; Length 468;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
 Db 401 GAVVPN 406
 Query Match 100.0%; Score 31; DB 10; Length 406;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
 ID 004236 PRELIMINARY; PRT; 72 AA.
 AC 004236;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
 DE Potassium channel (Fragment).
 OS Vicia faba (Broad bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eufrosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
 NCBI_TaxID=3906;
 RX SEQUENCE FROM N.A.
 RA TISSUE=Leaf;
 RC MEDLINE=97272307; PubMed=9114073;
 RA Hoch S., Dreyer I., Dietrich P., Becker D., Mueller-Roeber B.,
 RA Hedrich R.;
 RT "Molecular basis of plant-specific acid activation of K⁺ uptake
 channels."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4806-4810(1997).
 RC TISSUE=Leaf;
 RA Philippart K.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y09749; CAA70896.1; -.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K-channel_pore.
 DR Pfam; PF00520; Ion_trans; 1.
 KW Ionic channel; Transmembrane.
 FT NON_TER 1
 FT NON_TER 72
 SQ SEQUENCE 72 AA; 8259 MW; D8B6A36D43C2F00F CRC64;
 QY 1 GAVVPN 6
 Db 5 GAVVPN 10
 Query Match 96.8%; Score 30; DB 10; Length 72;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 ID 004241 PRELIMINARY; PRT; 72 AA.
 AC 004241;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
 DE Potassium channel (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 NCBI_TaxID=4577;
 RX SEQUENCE FROM N.A.
 RA STRAIN=LS5131B; TISSUE=Coleoptile;
 RC MEDLINE=97272307; PubMed=9114073;
 RA Hoch S., Dreyer I., Dietrich P., Becker D., Mueller-Roeber B.,
 RA Hedrich R.;
 RT "Molecular basis of plant-specific acid activation of K⁺ uptake
 channels."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4806-4810(1997).
 RC TISSUE=Leaf;
 RA Philippart K.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y09749; CAA70896.1; -.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K-channel_pore.
 DR Pfam; PF00520; Ion_trans; 1.
 KW Ionic channel; Transmembrane.
 FT NON_TER 1
 FT NON_TER 72
 SQ SEQUENCE 72 AA; 8259 MW; D8B6A36D43C2F00F CRC64;
 QY 1 GAVVPN 6
 Db 5 GAVVPN 10
 Query Match 96.8%; Score 30; DB 10; Length 72;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RP SEQUENCE FROM N.A.
RC STRAIN-L55311B; TISSUE-Coleoptile;
RA Philipp K.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y09747; CAZ0894.1; -.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR Pfam: PF00520; Ion_trans.1.
KW Ionic channel; Transmembrane.
FT NON_TER 1 1
FT SEQUENCE 72 AA; 8283 MW; 56B84B6AFE3D5DED CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 10; Length 72;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
DB 5 GAVIPN 10

RESULT 6
O42199 PRELIMINARY; PRT; 108 AA.
AC O42199;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Beta-2 microglobulin precursor (Fragment).
GN B2M.
OS Ictalarus pricei.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=64534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98435523; PubMed=9745011;
RA Cristofelino M.F., Benedetto R., Antao A., Wilson M.R., Chinchar V.G.,
RA Miller N.W., Clem L.W., McConnell T.J.;
RT "beta2-microglobulin of ictaluriid catfishes.";
RL Immunogenetics 48:339-343(1998).
DR EMBL: AF016044; AAC64993.1; -.
DR HSSP: P01888; IBMG.
DR InterPro: IPR007110; Ig_1-like.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; Ig_c1; 1.
DR PROSITE: PS50835; Ig_LIKE; 1.
DR PROSITE: PS00290; Ig_MHC; 1.
KM Signal.
FT NON_TER 1 1
FT SIGNAL <1 11
FT CHAIN 12 >108
FT NON_TER 108 108
FT SEQUENCE 108 AA; 12163 MW; 507EEF13B0A2A9AB CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 13; Length 108;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
DB 54 GAVIPN 59

RESULT 7
O931A6 PRELIMINARY; PRT; 340 AA.
AC O931A6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative lacI-family transcriptional regulator.
GN RA0038 OR SMA0078.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Bariloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL: AE007198; AAK64696.1; -.
DR InterPro: IPR000843; HTH_Lact.
DR InterPro: IPR001761; Periplabp/Lact.
DR Pfam: PF00356; lact; 1.
DR Pfam: PF00352; Peripla_BP_1like; 1.
DR SMART: SM00354; HTH_Lact; 1.
DR PROSITE: PS00356; HTH_Lact_FAMILY; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 340 AA; 36655 MW; C5722E9EC7FA32C8 CRC64;

Query Match
Best Local Similarity 96.8%; Score 30; DB 16; Length 340;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVVPN 6
DB 72 GAVIPN 77

RESULT 8
O8XTH4 PRELIMINARY; PRT; 411 AA.
AC O8XTH4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative NADH oxidase-related oxidoreductase protein
DE (EC 1.-.-.-).
GN RSP0136 OR RS02984.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1000.
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Aitai M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissendbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646076; CAD17287.1; -.
DR InterPro: IPR001155; Oxidored_FMN.
DR Pfam: PF00724; oxidored_FMN; 1.
KW Oxidoreductase; Plasmid; Complete proteome.
SQ SEQUENCE 411 AA; 44798 MW; 88AF324B7A305FCB CRC64;

```

Query Match 96.8%; Score 30; DB 16; Length 411;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVPN 6
 ||:|
 DB 12 GAVPN 17

RESULT 9

ID 039279 PRELIMINARY; PRT; 587 AA.
 AC 039279;

DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Counterpart of HSV-1 gene UL25 and VZV gene 34.
 GN 36.
 OS Equine herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=10331;

RN SEQUENCE FROM N.A.
 RC STRAIN=NS80567;

RX MEDLINE=98264497; PubMed=9603335;
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-4."
 RL J. Gen. Virol. 79:1197-1203(1998).

RN SEQUENCE FROM N.A.
 RC STRAIN=NS80567;

RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030027; AAC59553.1; -
 DR InterPro; IPR002493; UL25.
 DR Pfam; PF01499; UL25; 1.

SO SEQUENCE 587 AA; 63615 MW; B8F2F91F4CC958 CRC64;

Query Match 96.8%; Score 30; DB 12; Length 587;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVPN 6
 ||:|
 DB 413 GAVPN 418

RESULT 10

ID 08B9C8 PRELIMINARY; PRT; 803 AA.
 AC 08B9C8;

DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Rachioplusia ou multiple nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=80366;

RN SEQUENCE FROM N.A.
 RA Bonning B.C., Harrison R.L.;
 RT "The Rachioplusia ou multiple nucleopolyhedrovirus genome sequence."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY15471; AAN28018.1; -
 KM Hypothetical protein.

SO SEQUENCE 803 AA; 94545 MW; 30D414D7B37EF22E CRC64;

Query Match 96.8%; Score 30; DB 12; Length 803;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVPN 6
 ||:|
 DB 412 GAVPN 417

RESULT 11

ID 024538 PRELIMINARY; PRT; 807 AA.
 AC 024538;

DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Potassium channel (fragment).
 OS Vicia faba (Broad bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids; I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.

OX NCBI_TaxID=3906;

RN SEQUENCE FROM N.A.
 RC TISSUE=Colylodon;

RA Ache P., Mohlfarth T., Hedrich R., Becker D.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y10579; CAA71598.1; -

DR InterPro; IPR002110; ANK.
 DR InterPro; IPR005955; CNMP_binding.

DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001622; K-channel_pore.

DR InterPro; IPR005820; M-channel_nlg.
 DR Pfam; PF000023; ank; 5.

DR Pfam; PF000027; CNMP_binding; 1;
 DR Pfam; PF00520; Ion_trans; 1.

DR PRINTS; PRO1415; ANKYRIN.
 DR SMART; SM00248; ANK; 3.

DR SMART; SM00100; CNMP; 1.
 DR PROSITE; PSS0088; ANK_REPEAT; 1.

DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PSS0042; CNMP_BINDING_3; 1.

KW ANK repeat; Ionic channel; Repeat; Transmembrane.
 FT NON_TER 807 807

SO SEQUENCE 807 AA; 92464 MW; D0B01A2CD25A1038 CRC64;

Query Match 96.8%; Score 30; DB 10; Length 807;
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAVPN 6
 ||:|
 DB 230 GAVPN 235

RESULT 12

ID 09FY04 PRELIMINARY; PRT; 830 AA.
 AC 09FY04;

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Potassium channel 2.
 GN PTK2.

OS Populus tremula x Populus tremuloides.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids; I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=47664;

RN SEQUENCE FROM N.A.
 RC TISSUE=Phloem, Phloem, and Xylem;

RA Ache P., Langer K., Regan S., Geiger D., Hedrich R.;
 RT "Potassium dependent wood formation."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ271447; CAC05489.1; -
 InterPro; IPR002110; ANK.

DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR Pfam: PF00023; ank; 5.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR PRINTS: PR01415; ANKYRIN.
 DR SMART: SM00248; ANK; 3.
 DR SMART: SM00100; CNMP; 1.
 DR PROSITE: PSS0088; ANK_REPEAT; 2.
 DR PROSITE: PSS0297; ANK_REPEAT; 1.
 DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
 DR ANK repeat; Ionic channel; Repeat; Transmembrane.
 KW ANK repeat; Ionic channel; Repeat; Transmembrane.
 SQ SEQUENCE 830 AA; 94186 MW; 26D0458F845HC821 CRC64;

Query Match 96.8%; Score 30; DB 10; Length 830;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAVPN 6
 |||:|
 Db 237 GAVIPN 242

RESULT 13
 092PL4 PRELIMINARY; PRT: 832 AA.
 AC 092PL4;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Pulvinus inward-rectifying channel for potassium SPICK1.
 OS Samanea saman.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustoids 1; Fabales; Fabaceae; Mimosoidae; Ingeae; Samanea.
 OX NCBI_TaxID=76910;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moshelion M., Moran N., Hedrich R., Becker D.;
 RT "Molecular and biophysical analysis of Samanea saman motor cells.";
 RL Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF090905; AAD16278.1; -
 DR HSSP: Q54397; 1BL8.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR Pfam: PF00023; ank; 5.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR PRINTS: PR01415; ANKYRIN.
 DR SMART: SM00248; ANK; 2.
 DR SMART: SM00100; CNMP; 1.
 DR PROSITE: PSS0088; ANK_REPEAT; 3.
 DR PROSITE: PSS0297; ANK_REPEAT; 1.
 DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
 DR ANK repeat; Ionic channel; Repeat; Transmembrane.
 KW ANK repeat; Ionic channel; Repeat; Transmembrane.
 SQ SEQUENCE 832 AA; 95189 MW; EEF90E2A51CE4707 CRC64;

Query Match 96.8%; Score 30; DB 10; Length 832;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAVPN 6
 |||:|
 Db 255 GAVIPN 260

RESULT 14
 09SM12

ID 09SM12 PRELIMINARY; PRT: 849 AA.
 AC 09SM12;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Potassium channel protein ZMK2.
 GN ZMK2.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=L551311B, and cv. Apache; TISSUE=Coleoptile;
 RX MEDLINE=20604478; PubMed=10518597;
 RA Philippat K., Fuchs I., Luthen H., Roth S., Bauer C.S., Hage K.,
 Thiel G., Ljung K., Sandberg G., Bottger M., Becker D., Hedrich R.;
 RT "Auxin-induced K+ channel expression represents an essential step in
 coleoptile growth and gravitropism.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:12186-12191(1999).
 DR EMBL: AJ132686; CAB54856.1; -
 DR HSSP: P42773; 1IIB.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR Pfam: PF00023; ank; 4.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR SMART: SM00248; ANK; 2.
 DR SMART: SM00100; CNMP; 1.
 DR PROSITE: PSS0088; ANK_REPEAT; 3.
 DR PROSITE: PSS0297; ANK_REPEAT; 1.
 DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
 DR ANK repeat; Ionic channel; Repeat; Transmembrane.
 SQ SEQUENCE 849 AA; 93951 MW; D62EBA1276A2B07B CRC64;

Query Match 96.8%; Score 30; DB 10; Length 849;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAVPN 6
 |||:|
 Db 244 GAVIPN 249

RESULT 15
 097017 PRELIMINARY; PRT: 1891 AA.
 AC 097017;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE L-type calcium channel alpha-1 subunit.
 GN STPCACHL.
 OS Stylophora pistillata.
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
 OC Asterozoa; Pocilloporidae; Stylophora.
 OX NCBI_TaxID=50429;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99148007; PubMed=10023047;
 RA Zoccola D., Tambutte E., Senegas-Balas F., Michiels J.F., Failla J.P.,
 Jaubert J., Allemand D.;
 RT "Cloning of a calcium channel alpha1 subunit from the reef-building
 coral, Stylophora pistillata.";
 RL Gene 227:157-167(1999).
 DR EMBL: U64465; AAD11470.1; -
 DR InterPro: IPR001682; Ca/Na_pore.
 DR InterPro: IPR002111; Ca_channel_TripL.
 DR InterPro: IPR002077; Ca_channel.

DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR005446; LVDCALpha1.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00520; Ion_trans; 4.
 DR PRINTS; PR00167; CACHANNEL.
 DR PRINTS; PR01630; LVDCALPHA1.
 KW Ionic channel; Transmembrane.
 SO SEQUENCE 1891 AA; 213240 MW; BAB705AEBF9B4134 CRC64;

Query Match 96.8%; Score 30; DB 5; Length 1891;
 Best Local Similarity 83.3%; Pred. No. 7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAVYPN 6
 11:111
 Db 744 GAVYPN 749

Search completed: September 4, 2003, 21:06:40
 Job time : 24.7086 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 30.8824 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893D-60

Perfect score: 39

Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	6	23	AAE16725
2	39	100.0	6	23	AAE16728
3	35	89.7	56	22	AAE16728
4	35	89.7	56	22	AAE16728
5	35	89.7	334	22	ABE58457
6	35	89.7	342	21	AAE18365
7	35	89.7	451	21	AAE18364
8	35	89.7	477	21	AAE18363
9	35	89.7	613	22	ABE2602
			394	23	ABJ04660

10	34	87.2	405	23	ABP40414
11	33	84.6	6	23	AAE16714
12	33	84.6	6	23	AAE16717
13	33	84.6	6	22	ABG19091
14	33	84.6	67	22	AAU42280
15	33	84.6	106	23	ABP69636
16	33	84.6	106	23	ABG32265
17	33	84.6	106	23	ABG32266
18	33	84.6	106	23	ABP65233
19	33	84.6	107	23	ABP42213
20	33	84.6	111	23	ABP42647
21	33	84.6	127	23	ABP64899
22	33	84.6	131	21	AAE43900
23	33	84.6	137	23	ABP41758
24	33	84.6	138	22	ABG16314
25	33	84.6	144	22	AAE88390
26	33	84.6	147	22	ABG27187
27	33	84.6	200	17	AAE94701
28	33	84.6	200	18	AAE25953
29	33	84.6	320	19	AAE70220
30	33	84.6	320	19	AAE70236
31	33	84.6	320	23	AAE71270
32	33	84.6	320	23	AAE71286
33	33	84.6	320	23	AAE24934
34	33	84.6	320	23	AAE24950
35	33	84.6	320	23	ABG60872
36	33	84.6	320	23	ABG60888
37	33	84.6	320	23	AAU71816
38	33	84.6	320	23	AAU71832
39	33	84.6	456	22	ABG08774
40	33	84.6	623	24	ABJ25843
41	33	84.6	700	24	ABE94704
42	33	84.6	709	23	ABJ26443
43	33	84.6	777	23	AAE71320
44	33	84.6	777	23	ABE55398
45	33	84.6	1048	22	AAU07143

ALIGNMENTS

RESULT 1
AAE16725
AAE16725 standard; peptide; 6 AA.

XX
AAE16725;
09-APR-2002 (first entry)

XX
Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #5.

XX
Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiac;
XX
vasculotropic; hypertension; skin disorder; dermatological;
XX
coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
XX
hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
XX
restenosis post-angioplasty; antiangiinal; anticoagulant; thrombolytic.

XX
Ligamentum nuchae.
XX
Synthetic.

XX
W0200191700-A2.

XX
06-DEC-2001.

XX
30-MAY-2001; 2001WO-US17384.

XX
30-MAY-2000; 2000US-0580110.

XX
30-MAY-2000; 2000US-0580156.

XX
30-MAY-2000; 2000US-0580893.

XX
30-MAY-2000; 2000US-0584001.

(CONN-) CONNECTIVE TISSUE IMAGING LLC.

Staphylococcus epi
Ligamentum nuchae
Ligamentum nuchae
Novel human diagno
Propionibacterium
Human polypeptide
Human ribosomal pr
Hypoxia-regulated
Human ovarian anti
Human protein seq
Human cancer assoc
Human ovarian anti
Novel human diagno
Human membrane or
Novel human diagno
PRRSV VR 2385 ORF-
ORF 5 protein of p
Leishmania antigen
Leishmania antigen
L. major LmSP9 an
L. major LmSP9 ex
Leishmania major L
Leishmania antigen
Leishmania antigen
Leishmania antigen
Novel human diagno
Aspergillus fumiga
Human protein sequ
Aspergillus fumiga
L. major LmSP9 an
Drosophila melanog
Chicken CRIM1 prot

PI Mlts TF, Sandberg LB, Jimenez F;
 XX WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 XX Claim 24; Page 21; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 39; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CVVPNC 6
 |||||
 Db 1 CVVPNC 6
 |||||
 RESULT 2
 AAE16728 standard; peptide; 6 AA.
 XX
 AC AAE16728;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #8.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; candidant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 OS
 XX
 FH Key
 FT Modified-site 1 Location/Qualifiers
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 6 with copper as a chelating agent"
 FT Modified-site 6
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 FT
 FT
 XX W0200191700-A2.
 PN
 XX
 PD 06-DEC-2001.
 PD
 XX 30-MAY-2001; 2001WO-US17384.
 PF
 XX 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR

PR 30-MAY-2000; 2000US-0584001.
 PA
 XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 XX Claim 24; Page 21; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived cyclic peptide.
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 39; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CVVPNC 6
 |||||
 Db 1 CVVPNC 6
 |||||
 RESULT 3
 AAM83563
 ID AAM83563 standard; Protein; 56 AA.
 XX
 AC AAM83563;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen SEQ ID NO:11156.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 OS
 XX
 PN W0200157182-A2.
 PN
 XX
 PD 09-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01354.
 PF
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234497.
PR 25-SEP-2000; 2000US-0234498.
PR 26-SEP-2000; 2000US-0234584.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239835.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM:
WPI: 2001-483426/52.
N-PSDB; AAK56344.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -

XX
PS Claim 11: SEQ ID NO 11156; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.

XX Sequence 56 AA;

Query Match 89.7%; Score 35; DB 22; Length 56;
Best Local Similarity 66.7%; Pred. No. 58;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY - 1 CAVPNC 6
|:||||
DB 33 CMIPNC 38

RESULT 4
ID ABB58457 standard; Protein: 334 AA.

XX ABB58457;
XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2163.

XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB: ABL02560.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 2163; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB160511), expressed DNA
XX sequences (AB161840-AB16175) and the encoded proteins
XX (AB161737-AB161722).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 334 AA;

Query Match 89.7%; Score 35; DB 22; Length 334;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVPNC 6
|:||||
DB 242 CAVPNC 247

RESULT 5

ID AAG18365 standard; Protein: 342 AA.

XX AAG18365;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 19740.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 04-MAY-1999; 99US-0132048.

XX 05-MAY-1999; 99US-0132407.

XX 06-MAY-1999; 99US-0132485.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134768.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136382.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147036.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155859.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 35; DB 21; Length 342;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6
 DB 159 CMIPNC 164

RESULT 6
AAG18364
ID AAG18364 standard; Protein; 451 AA.
XX
AC AAG18364;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 19739.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128844.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135359.
PR 25-MAY-1999; 99US-0135623.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150936.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 35; DB 21; Length 451;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 268 CMI PNC 273

RESULT 7
AAG18363
ID AAG18363 standard; Protein: 477 AA.
XX
AC AAG18363;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 19738.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
XX
PD 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0156559.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	21-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160747.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145276.	PR	23-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	23-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	05-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	06-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	23-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			

PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0156559.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	18-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160747.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	23-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 89.7%; Score 35; DB 21; Length 477;
 Best Local Similarity 66.7%; Pred. No. 4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6
 ID 294 CWPNC 299

RESULT 8
 ABB62602
 ID ABB62602 standard; Protein; 613 AA.
 XX ABB62602;
 AC ABB62602;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide seq ID NO 14598.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX

OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL06705.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 14598; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 613 AA;
 SQ
 Query Match 89.7%; Score 35; DB 22; Length 613;
 Best Local Similarity 83.3%; Pred. No. 5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPNC 6
 11111
 DB 3 CAVPNC 8
 RESULT 9
 ABJ04660
 ID ABJ04660 standard; Protein; 394 AA.
 XX
 AC ABJ04660;
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE Protein of NOVX 17 SEQ ID No 42.
 XX
 KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipemic;
 KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
 KW tranquiliser; neuroleptic; antidiabetic; antitumor; antinflammatory;
 KW anti-HIV; antiallergic; antirheumatic; antiarthritis; NOVX; diabetes;
 KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
 KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
 KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
 KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
 KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
 KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
 KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
 KW gene therapy.
 XX
 XX Unidentified.
 OS

PN WO200246409-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46586.
 XX
 PR 06-DEC-2000; 2000US-251660P.
 PR 12-DEC-2000; 2000US-255029P.
 PR 08-JAN-2001; 2001US-260326P.
 PR 24-JAN-2001; 2001US-263800P.
 PR 20-FEB-2001; 2001US-269942P.
 PR 24-APR-2001; 2001US-286183P.
 PR 20-AUG-2001; 2001US-313627P.
 PR 12-SEP-2001; 2001US-318712P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Li L, Patturajan M, Shinkets RA, Casman SJ, Malyankar UM;
 PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;
 PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA;
 PI Boldog FL, Colman SD, Eisen AJ, Liu X, Padigaru M, Spaderna SK;
 PI Zernusen BD;
 XX
 DR WPI: 2002-54774/58.
 DR N-PSDB; ABR05473.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing cancer, diabetes, obesity, dyslipidaemia, anorexia, and
 PT metabolic, neurodegenerative, immune and hematopoietic disorders -
 XX
 PS Claim 1; Page 156; 421pp; English.
 XX
 CC The invention relates to an isolated polypeptide, designated NOVX,
 CC comprising a sequence fully defined in the specification. The isolated
 CC protein, its encoding polynucleotide or an antibody created from the
 CC protein is useful in the manufacture of a medicament for treating a
 CC syndrome associated with a human disease, preferably a NOVX-associated
 CC disorder, or for treating or preventing a NOVX-associated disorder in a
 CC subject, preferably human. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are also useful
 CC for treating or preventing metabolic disorders, diabetes, obesity,
 CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
 CC disease, Parkinson's disorder, immune disorders, haematopoietic
 CC disorders, and various dyslipidaemias, metabolic disturbances associated
 CC with obesity, the metabolic syndrome X, wasting disorders associated with
 CC chronic diseases, and cancer. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are useful for
 CC treating or preventing neurological disorders such as epilepsy, stroke,
 CC mental disorders including mood, anxiety, schizophrenic disorders,
 CC disorders of vesicular transport such as cystic fibrosis, diabetes
 CC mellitus, goiter, gastrointestinal disorders including ulcerative
 CC colitis, other conditions associated with abnormal vesicle trafficking
 CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
 CC arthritis. A cell comprising the vector of the invention is useful for
 CC producing non-human transgenic animals. The polynucleotide of the
 CC invention can be used to treat disorders by gene therapy. This sequence
 CC represents one of the isolated NOVX proteins of the invention.
 XX
 SQ Sequence 394 AA;
 Query Match 87.2%; Score 34; DB 23; Length 394;
 Best Local Similarity 83.3%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPNC 6
 11111
 DB 124 CVPNC 129
 RESULT 10
 ABP40414
 ID ABP40414 standard; Protein; 405 AA.
 XX

AC ABP40414;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SHQ ID NO:5259.
 XX
 KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 XX
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI: 2002-381255/41.
 DR N-Psdb; ABN92959.
 XX
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections
 XX
 PS Disclosure; SEQ ID 5259; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP5124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC C.N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 SQ Sequence 405 AA;
 Query Match 87.2%; Score 34; DB 23; Length 405;
 Best Local Similarity 83.3%; Pred. No. 5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPNC 6
 DB 210 CVCPCNC 215

RESULT 11
 AAE16714
 ID AAE16714 standard; peptide; 6 AA.
 XX
 AC AAE16714;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VPO derived peptide #8.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW stenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.

OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..6
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 20; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and stenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VPO derived cyclic peptide.
 XX
 SQ Sequence 6 AA;
 Query Match 84.6%; Score 33; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPNC 6
 DB 1 CVPPOC 6

RESULT 12
 AAE16717
 ID AAE16717 standard; peptide; 6 AA.
 XX
 AC AAE16717;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VPO derived peptide #11.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW stenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 6 with copper as a chelating agent"
 FT Modified-site 6
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mitts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 20; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.
 XX
 SO Sequence 6 AA:
 Query Match 84.6%; Score 33; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPVNC 6
 DB 1 CVPVPC 6
 1 CVPVPC 6

RESULT 13

ABG19091
 ID ABG19091 standard; Protein; 60 AA.
 XX
 AC ABG19091;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19082.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX

PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HXSE-) HXSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PDB; AAS83278.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 49450; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABC00010-ABC3037 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 60 AA:
 Query Match 84.6%; Score 33; DB 22; Length 60;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPVNC 6
 DB 20 CVPVNC 25
 1 CVPVNC 25

RESULT 14

AAU42280
 ID AAU42280 standard; Protein; 67 AA.
 XX
 AC AAU42280;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #3176.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 9.35294 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-60

Perfect score: 39

Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR_76:*
2: pir1:*
3: pir2:*
4: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	97.4	135	2 JC2211	hypothetical 14.8K
2	38	97.4	135	2 JC2209	hypothetical 14.9K
3	36	92.3	303	2 DB3082	probable permease
4	36	92.3	442	1 UKPG	u-plasminogen acti
5	35	89.7	334	2 A48151	sperm tail protein
6	35	89.7	472	2 T47436	protein kinase-11k
7	35	89.7	505	2 C82216	probable fumurate
8	35	89.7	507	2 A83105	probable fumurate
9	34	87.2	712	2 T16338	hypothetical prote
10	33	84.6	106	1 R6H036	ribosomal protein
11	33	84.6	106	1 R6R36	ribosomal protein
12	33	84.6	365	2 F70988	hypothetical prote
13	33	84.6	418	2 F70940	hypothetical prote
14	33	84.6	442	2 T37733	hypothetical zinc
15	33	84.6	782	2 A61625	tenascin-like prot
16	33	84.6	819	2 DB5440	Cu2+-transporting
17	33	84.6	905	2 T38980	hypothetical prote
18	33	84.6	952	2 T03158	legumetin protein 6
19	33	84.6	2180	2 A47651	zinc-finger protei
20	32	82.1	295	2 T22039	hypothetical prote
21	32	82.1	328	2 B65086	hydrogenase (EC 1.
22	32	82.1	328	2 A85959	hydrogenase-2 smal
23	32	82.1	328	2 A91114	hydrogenase-2 smal
24	32	82.1	328	2 A80885	hydrogenase-2 smal
25	32	82.1	407	2 T51226	hypothetical prote
26	32	82.1	436	2 H87793	protein C27A12.8 f
27	32	82.1	439	2 T43813	translation elonga
28	32	82.1	491	2 F87793	protein C27A12.6 f
29	32	82.1	497	2 G87793	protein C27A12.7 f

30	32	82.1	574	1 S76132	protein kinase pkn
31	32	82.1	1476	2 A45773	ketich protein, lon
32	32	82.1	1912	2 T29088	vitellogenin I, por
33	31	79.5	118	2 I55515	dynein-like protei
34	31	79.5	151	2 C69179	hypothetical prote
35	31	79.5	339	2 S73485	heterocyst maturat
36	31	79.5	530	2 A54965	probable vesicular
37	31	79.5	556	2 H69279	adenine deaminase
38	31	79.5	756	2 C91266	biodegradative arg
39	31	79.5	756	2 H86106	biodegradative arg
40	31	79.5	756	2 S56346	arginine decarboxy
41	31	79.5	770	1 TWBYA2	transcription fact
42	31	79.5	1136	2 T30878	dynein heavy chain
43	31	79.5	1498	2 E86302	hypothetical prote
44	30	76.9	103	2 S70182	hypothetical prote
45	30	76.9	125	2 T01750	glucose-1-phosphat

ALIGNMENTS

RESULT 1

hypothetical 14.8K protein, LIM5 - trumpet lily (fragment)

C:Species: Lilium longiflorum (Trumpet Lily)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: JC2211

R:Kobayashi, T.; Kobayashi, E.; Sato, S.; Hotta, Y.; Miyajima, N.; Tanaka, A.; Tabata

DNA Res. 1, 15-26, 1994

A:Title: Characterization of cDNAs induced in meiotic prophase in lily microsporocyte

A:Reference number: PC2136; MUID:96051386; PMID:7584025

A:Accession: JC2211

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-135 <KOB>

A:Cross-references: DDBJ:D21811; NID:9431151; PIDN:BA04835.1; PID:9431152

Query Match

Best Local Similarity 97.4%; Score 38; DB 2; Length 135;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6

DB 130 CVPNC 135

RESULT 2

hypothetical 14.9K protein, LIM13 - trumpet lily (fragment)

C:Species: Lilium longiflorum (Trumpet Lily)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: JC2209

R:Kobayashi, T.; Kobayashi, E.; Sato, S.; Hotta, Y.; Miyajima, N.; Tanaka, A.; Tabata

DNA Res. 1, 15-26, 1994

A:Title: Characterization of cDNAs induced in meiotic prophase in lily microsporocyte

A:Reference number: PC2136; MUID:96051386; PMID:7584025

A:Accession: JC2209

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-135 <KOB>

A:Cross-references: GB:D21819; NID:9442467; PIDN:BA04843.1; PID:9452595

Query Match

Best Local Similarity 97.4%; Score 38; DB 2; Length 135;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6

DB 130 CVPNC 135

RESULT 3

DB3082

probable permease of ABC transporter PA4504 [imported] - Pseudomonas aeruginosa (strain C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83082
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Muzoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
., Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83082
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <STO>
A:Cross-references: GB:AE004864; GB:AE004091; NID:99950740; PIDN:AA07892.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4504
C:Superfamily: oligopeptide permease protein oppB

Query Match 92.3%; Score 36; DB 1; Length 303;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
11:1111
Db 211 CVPNC 216

RESULT 4

UKRG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N:Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C:Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carboxylate (Asn) (covalent) #status predicted
F:179-310, 220-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 92.3%; Score 36; DB 1; Length 442;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
11:1111
Db 148 CVPNC 153

RESULT 5

A48151

sperm tail protein Mst98Ca - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
C:Accession: A48151; S23903
R:Schafer, M.; Borsch, D.; Hulster, A.; Schafer, U.
Mol. Cell. Biol. 13, 1708-1718, 1993
A:Title: Expression of a gene duplication encoding conserved sperm tail proteins is t
A:Reference number: A48151; MUID:93180818; PMID:8441407
A:Accession: A48151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <SCH>
A:Cross-references: EMBL:X67704; NID:98267; PID:98268
A:Note: sequence extracted from NCBI backbone (NCBIF:125984)
C:Genetics:
A:Gene: Mst98Ca
A:Cross-references: FlyBase:FBgn0002865

Query Match 89.7%; Score 35; DB 2; Length 334;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
11:1111
Db 242 CVPNC 247

RESULT 6

T47436
protein kinase-like protein - Arabidopsis thaliana
N:Alternate names: protein T18B22.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C:Accession: T47436
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224467
A:Accession: T47436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <JOR>
A:Cross-references: EMBL:AL138652
A:Experimental source: cultivar Columbia; BAC clone T18B22
C:Genetics:
A:Map position: 3
A:Introns: 182/2; 327/1
A:Note: T18B22.10
C:Superfamily: probable serine/threonine-specific protein kinase ATPK64; protein kina

Query Match 89.7%; Score 35; DB 2; Length 472;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
11:1111
Db 268 CVPNC 273

RESULT 7

C82216
probable fumarate hydratase, class I VC1304 [imported] - Vibrio cholerae (strain N169
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82216
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
charson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: C82216
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003952; NID:g9655789; PIDN:AAE94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: iron-dependent fumarate hydratase; iron-dependent tartrate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 272 CMIPNC 277

RESULT 8
A83105
Probable fumarate PA4333 [Imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83105
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <STO>
A:Cross-references: GB:AE004849; GB:AE004091; NID:g9950550; PIDN:AG07721.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4333
C:Superfamily: iron-dependent fumarate hydratase; iron-dependent tartrate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 507;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 272 CMIPNC 277

RESULT 9
T16338
hypothetical protein F42C5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
R:Du, Z.
Submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F42C5.
A:Reference number: Z18497
A:Accession: T16338
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-712 <DU2>
A:Cross-references: EMBL:U00799; NID:g1065935; PID:g1065937; PIDN:AAA81480.1; CESP:F42C5
C:Genetics:
A:Gene: CESP:F42C5.4
A:Introns: 20/1; 43/1; 106/3; 171/1; 391/3; 418/1; 455/3; 493/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F42C5.4

Query Match 87.2%; Score 34; DB 2; Length 712;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6

Db 425 CVPNC 430

RESULT 10
R6H036
ribosomal protein L36a - human
N:Alternate names: ribosomal protein HL44
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C:Accession: A25560
R:Davies, M.S.; Henney, A.; Ward, W.H.J.; Craig, R.K.
Gene 45, 183-191, 1986
A:Title: Characterisation of an mRNA encoding a human ribosomal protein homologous to
A:Reference number: A25560; MUID:87106812; PMID:3542712
A:Accession: A25560
A:Molecule type: mRNA
A:Residues: 1-106 <DAV>
A:Cross-references: EMBL:M15661; NID:g337577; PIDN:AAA36589.1; PID:g337578
A:Note: the authors translated the codon GGC for residue 29 as Glu
C:Genetics:
A:Gene: GDB:RPL36A
A:Cross-references: GDB:128850; OMIM:180469
A:Map position: 14pter-14qter
C:Superfamily: rat ribosomal protein L36a
C:Keywords: protein biosynthesis; ribosome
F:2-106/Product: ribosomal protein L36a #status predicted <MAT>

Query Match 84.6%; Score 33; DB 1; Length 106;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 72 CVPNC 77

RESULT 11
R6RT36
ribosomal protein L36a, cytosolic [validated] - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: A29820; A34987
R:Gallagher, M.J.; Chan, Y.L.; Lin, A.; Wool, I.G.
DNA 7, 269-273, 1988
A:Title: Primary structure of rat ribosomal protein L36a.
A:Reference number: A29820; MUID:88283346; PMID:3396452
A:Accession: A29820
A:Molecule type: mRNA
A:Residues: 1-106 <GAL>
A:Cross-references: EMBL:M16635; NID:g206731; PIDN:AAB54277.1; PID:g206732
A:Accession: A34987
A:Molecule type: protein
A:Residues: 2-6, 'X', 8-9, 'X', 11 <GAL2>
A:Note: the protein is designated as ribosomal protein L36a
C:Superfamily: rat ribosomal protein L36a
C:Keywords: protein biosynthesis; ribosome
F:2-106/Product: ribosomal protein L36a #status experimental <MAT>

Query Match 84.6%; Score 33; DB 1; Length 106;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 72 CVPNC 77

RESULT 12
F70988
hypothetical protein Rv1765c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 15-Sep-2000

C:Accession: F70988
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70988
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-365 <COL>
 A:Cross-references: GB:295890; GB:AL123456; NID:93242245; PIDN:CAB09341.1; PID:el300741;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV1765c
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1765c

Query Match 84.6%; Score 33; DB 2; Length 365;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 |||||
 Db 309 CVPNC 314

RESULT 13

F70940
 hypothetical protein RV2015c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 15-Sep-2000
 C:Accession: F70940
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70940
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-418 <COL>
 A:Cross-references: GB:AL021899; GB:AL123456; NID:93242282; PIDN:CAA17229.1; PID:el25199
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV2015c
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1765c

Query Match 84.6%; Score 33; DB 2; Length 418;
 Best Local Similarity 83.3%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 |||||
 Db 309 CVPNC 314

RESULT 14

T37733
 hypothetical zinc finger protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37733
 R:David, P.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21741
 A:Accession: T37733
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-442 <DAV>
 A:Cross-references: EMBL:AL121745; PIDN:CAB57403.1; GSPDB:GN00066; SPDB:SPAC16.05c

A:Experimental source: strain 97Zh-; cosmid c16
 C:Genetics:
 A:Gene: SPDB:SPAC16.05c
 A:Map position: 1

Query Match 84.6%; Score 33; DB 2; Length 442;
 Best Local Similarity 83.3%; Pred. No. 97;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 |||||
 Db 352 CVPNC 357

RESULT 15

A61625
 tenascin-like protein precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
 C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 11-Jan-2000
 C:Accession: A61625; S28463
 R:Baumgartner, S.; Chiquet-Ehrismann, R.; Mech. Dev. 40, 165-176, 1993
 A:Title: Ten(a), a Drosophila gene related to tenascin, shows selective transcript 10
 A:Reference number: A61625; MUID:93264270; PMID:7684246
 A:Accession: A61625
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-782 <BAU>
 A:Cross-references: EMBL:X68794
 C:Genetics:
 A:Gene: ten-a
 A:Cross-references: FlyBase:FBgn0004446
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: tandem repeat
 F:1-62/Domain: signal sequence #status predicted <SIG>
 F:63-782/Product: tenascin-like protein #status predicted <MAP>
 F:497-524/Domain: EGF homology <EGF>

Query Match 84.6%; Score 33; DB 2; Length 782;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 |||||
 Db 497 CVPNC 502

Search completed: September 4, 2003, 21:10:50
 Job time : 11.3529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 : Search time 5.20588 Seconds

(Without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893d-60

Perfect score: 39

Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	36	92.3	442	1	UROK_PIG
2	33	84.6	105	1	P09896 homo sapien
3	33	84.6	819	1	AHM1_ARATH
4	33	84.6	905	1	YAG1_SCHPO
5	33	84.6	1969	1	2292_HUMAN
6	32	82.1	328	1	HYBA_ECOLI
7	32	82.1	439	1	IFZG_ENCCU
8	32	82.1	574	1	SPKB_SYNY3
9	32	82.1	1477	1	KEIC_DROME
10	32	82.1	1912	1	VIT1_CHICK
11	31	79.5	339	1	4671_MYCPN
12	31	79.5	344	1	2DHA_HUMAN
13	31	79.5	423	1	WR58_ARATH
14	31	79.5	506	1	GAE_HUMAN
15	31	79.5	556	1	ADEC_ARCFU
16	31	79.5	652	1	P52K_MOUSE
17	31	79.5	755	1	ADIA_ECOLI
18	31	79.5	761	1	P52K_HUMAN
19	31	79.5	770	1	ACE2_YEAST
20	31	79.5	834	1	ANCI_CAEEL
21	30	76.9	125	1	GLGS_MAZE
22	30	76.9	146	1	VA28_YEAST
23	30	76.9	277	1	UBS1_YEAST
24	30	76.9	309	1	WR26_ARATH
25	30	76.9	311	1	Y467_MYCCE
26	30	76.9	332	1	PTA_CLOTM
27	30	76.9	398	1	ACTU_DROME
28	30	76.9	403	1	ARGD_BRUME
29	30	76.9	451	1	Y996_SYNY3
30	30	76.9	463	1	D5DR_FUGRU
31	30	76.9	473	1	GLGS_WHEAT
32	30	76.9	483	1	GLGS_ORYSA
33	30	76.9	489	1	GLGS_BETVU

34	30	76.9	506	1	GAE_RAT	Q9es14 ratus norv
35	30	76.9	508	1	GLGS_VICEA	P52416 vicia faba
36	30	76.9	512	1	GLGT_VICEA	P52417 vicia faba
37	30	76.9	512	1	WR33_ARATH	Q8s8p5 arabidopsis
38	30	76.9	513	1	GLGS_HORVU	P55238 hordeum vul
39	30	76.9	520	1	GLGS_ARATH	P55238 arabidopsis
40	30	76.9	520	1	GLGS_BRANA	Q9462 brassica na
41	30	76.9	521	1	GLGS_LYCES	Q42882 lycopersico
42	30	76.9	521	1	GLGS_SOLTU	P23509 solanum tub
43	30	76.9	525	1	ACU8_NEUCR	P13937 neurospora
44	30	76.9	606	1	SP2_HUMAN	Q02086 homo sapien
45	30	76.9	644	1	BTD_DROME	Q24266 drosophila

ALIGNMENTS

```

RESULT 1
UROK_PIG          STANDARD;          PRT;          442 AA.
AC  P04185:
DT  20-MAR-1987 (Rel. 04, Created)
DT  13-AUG-1987 (Rel. 05, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE  (U-plasminogen activator).
GN  PLAU.
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX  NCBI_TaxID=9823;
RN  [1]
RP  TISSUE: FROM N.A.
RC  TISSUE: Kidney;
RX  MEDLINE=85087954; PubMed=6096832;
RA  Negamine Y., Pearson D., Altus M.S., Reich E.;
RT  "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL  Nucleic Acids Res. 12:9525-9541(1984).
RN  [2]
RP  REVISION TO 241.
RA  Negamine Y.;
RL  Submitted (DEC-1986) to the PIR data bank.
CC  -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC  plasminogen to form plasmin.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC  -1- SIMILARITY: Contains 1 kringle domain.
CC  -1- SIMILARITY: Contains 1 EGF-like domain.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL: X01648; CAA25806.1; -
EMBL: X02724; CAA26511.1; -
PIR: A00932; UKRG.
HSSP: P00749; IKDU.
DR  MEROPS: S01.231; -.
DR  InterPro: IPR001314; Chymotrypsin.
DR  InterPro: IPR006209; EGF like.
DR  InterPro: IPR000001; Kringle.
DR  InterPro: IPR001254; Ser-protease-Try.
DR  Pfam: PF00051; Kringle; 1.
DR  Pfam: PF00089; trypsin; 1.
DR  PRINTS: PR00722; CHYMOTRYPSIN.
DR  PRINTS: PR00018; KRINGLE.
DR  PRODOM: PD000395; Kringle; 1.
DR  SMART: SM00130; KR; 1.
DR  SMART: SM00020; TRYPSIN; 1.
DR  PROSITE: PS00022; EGF_1; 1.

```

DR PROSITE: PS01186; EGF_2; FALSE-NEG.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin.DOM; 1.
 DR PROSITE: PS00134; TRYPsin.HIS; 1.
 DR PROSITE: PS00135; TRYPsin.SER; 1.
 KM Plasmogen activation; Hydrilase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 442
 FT CHAIN 21 188
 FT CHAIN 190 442
 FT CHAIN 29 65
 FT CHAIN 72 153
 FT CHAIN 154 189
 FT CHAIN 190 442
 FT CARBOHYD 152 152
 FT DISULFID 33 41
 FT DISULFID 35 53
 FT DISULFID 55 64
 FT DISULFID 179 310
 FT DISULFID 220 236
 FT DISULFID 228 299
 FT DISULFID 324 393
 FT DISULFID 356 372
 FT DISULFID 383 411
 FT ACT_SITE 235 235
 FT ACT_SITE 286 286
 FT ACT_SITE 387 387
 FT CONFLICT 241 241
 FT CONFLICT 242 242
 FT CONFLICT 288 288
 SQ SEQUENCE 442 AA: 49116 MW: 83364
 Best Local Similarity 92.3%; Score 36; DB 1; Length 442;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
 DB 148 CVPNC 153
 ID 144 HUMAN STANDARD: PRT: 105 AA.
 AC P09896: P10661:
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 60S ribosomal protein L44 (L36a).
 GN RPL44 OR RPL36A.
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat), and
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606, 10090, 10116, 9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human;
 RX MEDLINE=87106812; PubMed=3542712;
 RA Davies M.S., Henney A., Ward W.H.J., Craig R.K.;
 RT "Characterisation of an mRNA encoding a human ribosomal protein
 homologous to the yeast L44 ribosomal protein.";
 RL Gene 45:193-191(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human;
 RA Oeltjen J.C., Malley T.M., Muzny D.M., Miller W., Gibbs R.A.,
 RA Belmont J.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human; TISSUE-Eye;
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse; STRAIN-C129;
 RX MEDLINE=95352959; PubMed=7626884;
 RA Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D., Belmont J.W.,
 RA Gibbs R.A.;
 RT "Sixty-nine kilobases of contiguous human genomic sequence containing
 the alpha-galactosidase A and Brulon's tyrosine kinase loci.";
 RL Mamm. Genome 6:334-338(1995).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse; STRAIN-C57BL/6J; TISSUE-Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinozawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse; TISSUE-Mammary gland;
 RX MEDLINE=22388257; PubMed=1247932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Merra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC SPECIES-Rat;
 RX MEDLINE=88283346; PubMed=3396452;
 RA Gallagher M.J., Chan Y.-L., Lin A., Wool I.G.;
 RT "Primary structure of rat ribosomal protein L36a.";
 RL DNA 7:269-273(1988).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Pig;
 RA Kimura M., Kawakami K., Suzuki H., Hamasima N.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: Cytoplasmic.
 CC - SIMILARITY: BELONGS TO THE L44E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15661; AAA36589.1; -
 DR EMBL: U78027; AAB64204.1; -
 DR EMBL: BC001781; AAH01781.1; -
 DR EMBL: U58105; AAB47245.1; -
 DR EMBL: AK002540; BAB22175.1; -
 DR EMBL: AK003166; BAB22616.1; -
 DR EMBL: AK010629; BAB27075.1; -
 DR EMBL: AK012502; BAB28285.1; -
 DR EMBL: BC019810; AAH19810.1; -
 DR EMBL: BC027515; AAH27515.1; -
 DR EMBL: M19635; AAB54277.1; -
 DR EMBL: AB000910; BAA19210.1; -
 DR PIR: A29820; K6RT36.
 DR Gene: HGNC:10359; RPL36A.
 DR GK: P09896; -
 DR MIM: 180469; -
 DR MGD: MGI:1201789; RPL14.
 DR GO: GO:0005840; Cytoplasm; NAS.
 DR InterPro: IPR000552; Ribosomal_L44E.
 DR Pfam: PF00935; Ribosomal_L44; 1.
 DR ProDom: PD002841; Ribosomal_L44E; 1.
 DR ProSITE: PS01172; RIBOSOMAL_L44E; 1.
 DR KMW Ribosomal protein.
 FT INIT MET 0
 FT CONFLICT 37 37 K -> R (IN REF. 1).
 FT SEQUENCE 105 AA; 12310 MW; 4C599DD5924C4EB3 CRC64;
 Query Match 84.6%; Score 33; DB 1; Length 105;
 Best Local Similarity 83.3%; Pred. No. 5;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CWPNC 6
 Db 71 CWPNC 76
 RESULT 3
 AHM1_ARATH STANDARD; PRT; 819 AA.
 ID AHM1_ARATH
 AC Q9M3H5; Q9SMW6;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Potential cadmium/zinc-transporting ATPase HMA1 (EC 3.6.3.3)
 DE (EC 3.6.3.5).
 GN HMA1 OR AT4G37270 OR C7A10.90 OR AP22.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA MEDLINE=20083488; PUBMed=10617198;
 RA Mayer K.F.X., Scheller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansgore W., Brandt P., Grievell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delseny M., Putgdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hohisel J., Zimmermann W., Wedler H., Riddle P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chang Y.-J., Vandebussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Mooliman P., Klein Lankhorst R., Rose M., Haut J., Koetter P.,
 RA Bernerster S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 RA Petzelt A., Rajandream M.A., Lyne M., Benes V., Reichenart S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Farman B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argilou A., Vitale D., Liguori R., Pirvadi E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schmal S., Hiller R., Schmidt W., Lecharry A., Aubourg S.,
 RA Cheifor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heljnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kallio J., Graves T., Harmon G., Edwards J.,
 RA Latrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton L., Mardis E., Dante M., Pepin K., Hiller L.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hiller L.,
 RA Nelson J., Spith J., Ryan E., Andrews S., Gelsel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Grant S., Shohdy N., Hasegawa A., Hamed A., Lohli M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana."
 RL Nature 402:769-777(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RA Page S.L., Pittman J.K., Krijger G.C., Williams L.E.;
 RT "Identification of a putative heavy metal P-type ATPase in
 Arabidopsis."
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: INVOLVED IN CADMIUM/ZINC TRANSPORT (POTENTIAL).
 CC - CATALYTIC ACTIVITY: ATP + H(2)O + Cd(2+)(in) = ADP + phosphate +
 Cd(2+)(out).
 CC - CATALYTIC ACTIVITY: ATP + H(2)O + Zn(2+)(in) = ADP + phosphate +
 Zn(2+)(out).
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: Belongs to the cation transport ATPases family (P-type
 ATPases). Subfamily IB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z99707; CAB16773.1; -
 DR EMBL: AL161591; CAB80393.1; -
 DR EMBL: AJ400906; CAB90352.1; -
 DR PIR: D85440; D85440.
 DR InterPro: IPR006416; ATPase-IB_hvy.
 DR InterPro: IPR001757; ATPase-IB_hvy.
 DR InterPro: IPR006404; Heavy_metal_ATPase.
 DR InterPro: IPR005834; Hydrolase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.

```

DR Pfam: PF00702; Hydrolase: 1.
DR PRINTS; PRO0119; CATAPSE.
DR TIGRfams; TIGR01512; ATPase-IB2_Cd; 1.
DR TIGRfams; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRfams; TIGR01494; ATPase-P-type; 2.
DR PROSITE; PS00154; ATPase_E1_E2; 1.
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding; Magnesium;
KW Zinc; Cadmium; Metal-binding; Multigene family.
FT DOMAIN 1 122
FT TRANSMEM 123 144
FT DOMAIN 145 153
FT TRANSMEM 154 173
FT DOMAIN 174 180
FT TRANSMEM 181 201
FT DOMAIN 202 202
FT TRANSMEM 203 223
FT DOMAIN 224 361
FT TRANSMEM 362 384
FT TRANSMEM 385 398
FT TRANSMEM 399 416
FT TRANSMEM 417 737
FT TRANSMEM 738 757
FT TRANSMEM 758 762
FT TRANSMEM 782 819
FT DOMAIN 819 89
FT TRANSMEM 809 817
FT MOD_RES 453 453
FT METAL 662 682
FT METAL 686 686
FT VARIANT 53 53
FT VARIANT 105 105
FT VARIANT 659 659
SQ SEQUENCE 819 AA; 8818 MW; F281E8F3C0ED52 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 1; Length 819;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 485 CCIPNC 490

RESULT 4
YAG1_SCHPO STANDARD; PRT; 905 AA.
AC 009866; 09DH9;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical WD-repeat protein c12612.01c in chromosome 1.
GN SPAC12612.01c OR SPAC630.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetes; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
[1]
SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

```

```

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Wolckaert G., Aert R., Robben J., Gymnopoulos B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesli D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallade V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: Contains 4 WD repeats.
CC -1- SIMILARITY: TO YEAST YBL104C.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL109832; CAB52724.1; -
DR EMBL; Z66568; CA91496.1; -
DR PIR; T38980; T38980.
DR Genedb_Spombe; SPAC12612.01c; -
DR Pfam; PF00400; WD40; 2.
DR InterPro; IPR001680; WD40.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR PROSITE; PS50082; WD_REPEATS_2; FALSE_NEG.
DR PROSITE; PS50294; WD_REPEATS_REGION; FALSE_NEG.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 42 82
FT REPEAT 86 128
FT REPEAT 136 175
FT REPEAT 177 217
SQ SEQUENCE 905 AA; 101948 MW; 56BCF35CA07F709B CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 1; Length 905;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 893 CVPNC 898

RESULT 5
2292_HUMAN STANDARD; PRT; 1969 AA.
ID 2292_HUMAN
AC 060281; Q9H8J4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Zinc finger protein 292.
GN ZNF292 OR KIAA0530.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE OF 1-874 AND 1160-1786 FROM N.A.
RP TISSUE=Placenta;
RC Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

```

RA Yamamoto J., Makamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niino I., Iwanaga T.,
 RT "MEDO human cDNA sequencing project."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 407-1969 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AK023626; BAB14622.1; -
 DR EMBL: AK023712; BAB14654.1; -
 DR EMBL: AB011102; BAA23456.1; -
 DR HSSP: P08153; INCS.
 DR Genew: HGNC:18410; ZNF292.
 DR InterPro: IPR007087; Znf.C2H2.
 DR Pfam: PF00096; zf-C2H2; 17.
 DR SMART: SM00355; ZNF_C2H2; 12.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 9.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 7.
 KW Transcription regulation; zinc-finger; Metal-binding; Nuclear protein;
 KW DNA-binding; Repeat.
 FT ZN_FING 25 49 C2H2-TYPE.
 FT ZN_FING 53 77 C2H2-TYPE.
 FT ZN_FING 344 369 C2H2-TYPE.
 FT ZN_FING 1193 1219 C2H2-TYPE.
 FT ZN_FING 1360 1385 C2H2-TYPE.
 FT ZN_FING 1418 1443 C2H2-TYPE.
 FT ZN_FING 1462 1487 C2H2-TYPE.
 FT ZN_FING 1502 1527 C2H2-TYPE.
 FT ZN_FING 1632 1656 C2H2-TYPE.
 FT CONFLICT 601 601 E -> K (IN REF. 1).
 FT CONFLICT 776 776 N -> S (IN REF. 1).
 FT CONFLICT 1291 1291 I -> V (IN REF. 1).
 FT CONFLICT 1408 1408 E -> K (IN REF. 1).
 SO SEQUENCE 1969 AA; 218445 MW; 3b5d25d5adcdpasa CRC64;
 Query Match 84.6%; Score 33; DB 1; Length 1969;
 Best Local Similarity 83.3%; Pred. No. 86;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPNC 6
 DB 1195 CVPNC 1200
 ID HYBA_ECOLI STANDARD; PRT; 328 AA.
 AC P37179;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hydrogenase-2 operon protein hybA precursor.
 GN HYBA OR B2996 OR Z4350 OR ECS3881 OR SF3043.
 OS Escherichia coli,

OS Escherichia coli O157:H7, and
 OS Shigella flexneri.
 OC Bacteria: Proteobacteria: Gammaproteobacteria: Enterobacteriales;
 OC Enterobacteriaceae: Escherichia.
 ON NCBI_TaxID=562, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / TGI;
 RX MEDLINE=94292472; PubMed=8021226;
 RA Menon N.K., Chatelus C.Y., Dervartanian M., Wendt J.C.,
 RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.;
 RT "Cloning, sequencing, and mutational analysis of the hyb operon
 RT encoding Escherichia coli hydrogenase 2."
 RL J. Bacteriol. 176:4416-4423(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lim A., DiMantola E.T., Potamousis K.,
 RA Aoodca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kohara S., Shiba T., Hattori M., Shiga H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao Y., Gao Y., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.H7."
 RL Nucleic Acids Res. 30:4432-4441(2002).
 CC -1- FUNCTION: PARTICIPATES IN THE PERIPLASMIC ELECTRON-TRANSFERRING
 CC ACTIVITY OF HYDROGENASE 2 DURING ITS CATALYTIC TURNOVER.
 CC -1- COFACTOR: BINDS 3 4FE-4S CLUSTERS AND A 3FE-4S CLUSTER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC BACTERIAL-TYPE 4FE-4S FERREDOXINS.
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE THE SMALL SUBUNIT
 CC OF HYDROGENASE 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 009177; AAA21589.1; -
 CC EMBL; U28377; AAA69163.1; -
 CC EMBL; AE000382; AAC76032.1; -
 CC EMBL; AE005529; AAG58133.1; -
 CC EMBL; AP002563; BAB37304.1; -
 CC EMBL; AE015316; AAN44521.1; -
 CC PIR; A91114; A91114.
 CC PIR; B65086; B65086.
 CC Ecocyc; B811799; hyda.
 CC InterPro; IPR001450; 4fe4S_ferredoxin.
 CC InterPro; IPR006311; Tal.
 CC Pfam; PF000037; fer4; 1.
 CC TIGRfams; TIGR01409; Tat_signal_seq; 1.
 CC PROSITE; PS00198; 4FE4S_FERRDOXIN; 1.
 CC Oxidoreductase; Signal; Periplasmic; Iron-sulfur; 4fe-4S; 3fe-4S;
 KW Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 328
 FT METAL 47 478
 FT METAL 50 50
 FT METAL 53 53
 FT METAL 57 57
 FT METAL 112 112
 FT METAL 115 115
 FT METAL 120 120
 FT METAL 124 124
 FT METAL 145 145
 FT METAL 148 148
 FT METAL 151 151
 FT METAL 155 155
 FT METAL 174 174
 FT METAL 177 177
 FT METAL 193 193
 FT METAL 197 197
 SO SEQUENCE 328 AA; 36003 MW; 77203A0F50F61662 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 328;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 115 CVPNC 120

RESULT 7
 ID IF2G_ENCCU STANDARD; PRT; 439 AA.
 AC 096719;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Eukaryotic translation initiation factor 2 gamma subunit (eif-2-gamma).
 OS Eucephalitozoan cuniculi.
 CC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Eucephalitozoan.
 CC NCBI_Taxid=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Name:
 RA Duffieux F., Peyret P., Roe B.A., Vivares C.P.;
 RT "Putative elongation factor 2, from Eucephalitozoan cuniculi.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER

ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
 CC SUBUNIT: Heterotrimer composed of an alpha, a beta, and a gamma chain.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY. EIF2G SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ006823; CAN07260.1; -
 CC PIR; T43813; T43813.
 CC InterPro; IPR000795; EF_GTPbind.
 CC InterPro; IPR004161; EFTU_D2.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC PRINTS; PR00315; ELONGATNCT.
 KW Initiation factor; Protein biosynthesis; GTP-binding.
 FT NP_BIND 20 27
 FT NP_BIND 103 107
 FT NP_BIND 159 162
 SO SEQUENCE 439 AA; 47846 MW; 339B89D4FBDPCB CRC64;

Query Match 82.1%; Score 32; DB 1; Length 439;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 67 CVPNC 72

RESULT 8
 ID SPKB_SYNY3 STANDARD; PRT; 574 AA.
 AC P74297;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine-protein kinase B (BC 2.7.1.37).
 GN SPKB OR SLR1697.
 OS Synechocystis sp. (strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 CC NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamei A., Ikeuchi M.;
 RT "A novel gene, spkb, encodes active Ser/Thr protein kinase in the motile cyanobacterium Synechocystis sp. PCC 6803.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: Contains 2 pentapeptide repeat domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB046598; BAI17034.1; -
 DR EMBL: D90914; BAI18391.1; -
 DR PIR: S76132; S76132.
 DR InterPro: IPR001646; 5peptide_repeat.
 DR InterPro: IPR000719; Ser_Lhr_kinase.
 DR InterPro: IPR002290; Ser_Lhr_kinase.
 DR Pfam: PF00805; Pentapeptide; 2.
 DR Pfam: PF00699; Pentapeptide; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW Complete proteome.
 FT DOMAIN 34 301 PROTEIN KINASE.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 65 65 ATP (BY SIMILARITY).
 FT ACT_SITE 163 163 BY SIMILARITY.
 FT DOMAIN 454 493 PENTAPEPTIDE 1.
 FT DOMAIN 504 543 PENTAPEPTIDE 2.
 SQ SEQUENCE 574 AA; 63079 MW; D96B2C5ABEE405A3 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 574;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNCC 6
 11111
 Db 4 CVPNCC 9

RESULT 9
 KEEC_DROME STANDARD; PRT; 1477 AA.
 ID KEEC_DROME STANDARD; PRT; 1477 AA.
 AC 004652; 004653; 09YUZA;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ring canal kelch protein [contains: kelch short protein].
 GN KRL OR CG7210.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SOURCE FROM N.A.
 RC TISSUE-Embryo;
 RX MEDLINE=93201592; PubMed=84535663;
 RA Xue F., Cooley L.;
 RT "Kelch encodes a component of intercellular bridges in Drosophila egg
 RT chambers.";
 RL Cell 72:681-693(1993).
 RN [2]
 RP SOURCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bayandale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintinas S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burdick K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jalali M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., Moleda M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshfeil A.,
 RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP CHARACTERIZATION.
 RC TISSUE-Embryo;
 RX MEDLINE=97236487; PubMed=9118811;
 RA Robinson D.N., Cooley L.;
 RT "Examination of the function of two kelch proteins generated by stop
 RT codon suppression.";
 RL Development 124:1405-1417(1997).
 CC -1- FUNCTION: COMPONENT OF RING CANALS THAT REGULATES THE FLOW OF
 CC CYTOPLASM BETWEEN CELLS. MAY BE INVOLVED IN THE REGULATION OF
 CC CYTOPLASM FLOW FROM NURSE CELLS TO THE OOCYTE DURING OOGENESIS.
 CC BINDS ACTIN.
 CC -1- SUBCELLULAR LOCATION: INNER SURFACE OF CYTOPLASMIC BRIDGES OR RING
 CC CANALS PRESENT IN EGG CHAMBERS. SUBCORTICALLY IN IMAGINAL DISK
 CC EPITHELIA.
 CC -1- TISSUE SPECIFICITY: BOTH PROTEINS ARE EXPRESSED IN OVARIES. MALE
 CC TESTIS. OVARECTOMIZED FEMALES, CUTICLE, SALIVARY GLAND AND
 CC IMAGINAL DISKS. KELCH ORF1 IS THE PREDOMINANT PROTEIN AND IS ALSO
 CC EXPRESSED IN FAT BODIES. ON ENTRY INTO METAMORPHOSIS LEVELS OF
 CC FULL LENGTH PROTEIN INCREASE IN TESTIS AND IMAGINAL DISKS.
 CC -1- DEVELOPMENTAL STAGE: LARVAE, PUPAE AND ADULTS.
 CC -1- SIMILARITY: Contains 1 PBP/POZ domain.
 CC -1- SIMILARITY: Contains 6 kelch repeats.
 CC -1- CAUTION: REF.3 BELIEVES RESIDUE 690 IS A SILENCING PEPTIDE.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: L08483; AAA53471.1; -
 EMBL: L08483; AAA53472.2; -
 EMBL: AE003657; AAF53651.1; ALT-SEQ.
 HSSP: Q05516; ICS3.
 DR FlyBase: FBgn0001301; kel.
 DR GO: GO:0007292; P: oogenesis; IMP.
 DR GO: GO:0007301; P: ring canal formation; IDA.
 DR InterPro: IPR000210; BTB_POZ.
 DR InterPro: IPR006651; Kelch.
 DR InterPro: IPR006652; Kelch_rep.
 DR Pfam: PF00651; BTB; 1.

DR Pfam; PF01344; Kelch; 6.
 DR PRINTS; PR00501; KELCHREPEAT.
 DR SMART; SM00225; BTB; 1.
 DR SMART; SM00612; Kelch; 6.
 DR PROSITE; PS50097; BTB; 1.
 KW Cytoskeleton; Actin-binding; Selenium; Selenocysteine; Kelch repeat;
 KW Repeat.
 FT CHAIN 1 1477
 FT CHAIN 1 689
 FT DOMAIN 157 223
 FT REPEAT 404 449
 FT REPEAT 450 496
 FT REPEAT 498 543
 FT REPEAT 545 592
 FT REPEAT 594 639
 FT REPEAT 641 687
 FT DOMAIN 18 28
 FT DOMAIN 29 36
 FT DOMAIN 78 83
 FT SE_CYS 690 690
 FT CONFLICT 493 493
 FT CONFLICT 596 596
 FT CONFLICT 824 824
 FT CONFLICT 858 858
 FT CONFLICT 1083 1083
 FT CONFLICT 1086 1086
 SQ SEQUENCE 1477 AA; 160086 MW; 4851EEAE9D9DA47 CRC64;
 Query Match 82.1%; Score 32; DB 1; Length 1477;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 CVPNC 6
 Db 1025 CNVNC 1030
 RESULT 10
 VITL_CHICK STANDARD; PRT; 1912 AA.
 AC P87498;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vitellinogen I precursor (Minor vitellinogen) [Contains: Lipovitellin
 DE I (LVI); Phosvitin (PV); Lipovitellin II (LVI); YGP42].
 GN VITL OR VTGI.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OC NCBI_TaxID=9031;
 OX (1)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RA Mabuchi N., Yamamura J.-I., Adachi T., Aoki N., Nakamura R.,
 RA Matsuda T.,
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1628-1639, AND IDENTIFICATION OF YGP42.
 RX MEDLINE=95322425; PubMed=7599159;
 RA Yamamura J.-I., Adachi T., Aoki N., Nakajima H., Nakamura R.,
 RA Matsuda T.;
 RT "Precursor-product relationship between chicken vitellinogen and the
 RT yolk proteins: the 40 kDa yolk plasma glycoprotein is derived from
 RT the C-terminal cysteine-rich domain of vitellinogen II.";
 RL Biochim. Biophys. Acta 1244:384-394(1995).
 CC -I- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE
 CC SOURCES OF NUTRIENTS DURING EARLY DEVELOPMENT OF OVIPAROUS
 CC ORGANISMS.
 CC -I- FUNCTION: PHOSVITIN IS BELIEVED TO BE OF IMPORTANCE IN
 CC SEQUESTERING CALCIUM, IRON AND OTHER CATIONS FOR THE DEVELOPING

CC EMBRYO.
 CC -I- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD
 CC AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENDOCYTOSIS INTO GROWING
 CC OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE
 CC RESPECTIVE YOLK COMPONENTS.
 CC -I- INDUCTION: BY steroids (estrogen).
 CC -I- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST
 CC HIGHLY PHOSPHORYLATED (10%) PROTEINS IN NATURE.
 CC -I- PTM: CATHESPIN D IS RESPONSIBLE FOR INTRAOOCYTIC PROCESSING OF
 CC VITELLOGENIN.
 CC -I- PTM: MAY CONTAIN INTRAMOLECULAR DISULFIDE BONDS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D89547; BAA13973.1; -
 DR PIR; T29088; T29088.
 DR InterPro; IPR001747; Lipid_transprt_N.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF01347; Vitellinogen_N; 1.
 DR Pfam; PF00094; vwd; 1.
 DR SMART; SM00638; LPD_N; 1.
 DR SMART; SM00216; vwd; 1.
 KW Glycoprotein; Phosphorylation; Storage protein; Signal.
 FT CHAIN 1 15
 FT CHAIN 16 1912
 FT CHAIN 16 1139
 FT CHAIN 1140 1391
 FT CHAIN 1392 1627
 FT CHAIN 1628 1912
 FT DOMAIN 1093 1104
 FT DOMAIN 1150 1299
 FT DOMAIN 1359 1397
 FT DOMAIN 1612 1615
 FT CARBOHYD 163 163
 FT CARBOHYD 991 991
 FT CARBOHYD 1206 1206
 FT CARBOHYD 1375 1375
 FT CARBOHYD 1662 1662
 FT CARBOHYD 1698 1698
 FT CARBOHYD 1703 1703
 SQ SEQUENCE 1912 AA; 210630 MW; 89BA6273D6492E51 CRC64;
 Query Match 82.1%; Score 32; DB 1; Length 1912;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 CVPNC 6
 Db 791 CIVPTC 796
 RESULT 11
 Y467_MYCPN STANDARD; PRT; 339 AA.
 ID Y467_MYCPN
 AC P75110;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein MG467 homolog
 DE (K05_orf359).
 GN MP683 OR MP159.
 OS Mycoplasma pneumoniae.
 CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;

```

RX MEDLINE-97105885; PubMed-8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000017; AAB95807.1; -
DR PIR; S73485; S73485.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT NP_BIND 150 157 ATP (POTENTIAL).
SQ SEQUENCE 339 AA; 38138 MW; DAC0D8F08294E6AB CRC64;

Query Match 79.5%; Score 31; DB 1; Length 339;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 174 CVPANC 179

RESULT 12
ZDH4_HUMAN
ID ZDH4_HUMAN STANDARD; PRT; 344 AA.
AC O9NP88; O9NP88.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 4 (Zinc finger protein 374)
DE (DC1).
GN ZDHHC4 OR ZNF374.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Y., Peng Y., Li N., Gu W., Han Z., Fu G., Chen Z.;
RT "Novel genes expressed in human dendritic cell";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiraori A., Sudo H.,
RA Watanabe M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Nomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-21154917; PubMed-11230166;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,

```

```

RA Ansoorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duetschoelt A., Beyer A., Koehner K., Strack N.,
RA Mewes H.-W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs";
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RX MEDLINE-22388257; PubMed-12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marsina K., Parner A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Yoshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huljk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellton E., Kelleman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Katta W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK001341; BAA91636.1; -
DR EMBL; AF201931; AAF86867.1; -
DR EMBL; AL136674; CAB66609.1; -
DR EMBL; BC001239; AAH01239.1; -
DR Genew; HGNC:18471; ZDHHC4.
DR InterPro; IPR001594; ZnF_DHHC.
DR Pfam; PF01529; zf-DHHC; 1.
DR ProDom; PD003041; ZnF_DHHC; 1.
DR PROSITE; PS50216; ZF_DHHC; 1.
KW Transmembrane, Zinc-finger.
FT TRANSMEM 3 23
FT TRANSMEM 68 88
FT TRANSMEM 100 120
FT TRANSMEM 193 213
FT TRANSMEM 234 254
FT TRANSMEM 256 276
FT ZN_FING 149 199
FT TRANSMEM 244 244
FT TRANSMEM 276 276
FT TRANSMEM 294 294
FT TRANSMEM 344 344
FT TRANSMEM 39786 MW; 3D2CF22B14036908 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 344;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 43 CVPNC 48

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U06661; AAB49284.1; -
 DR EMBL: Y09763; CAA70903.1; -
 DR EMBL: Y09764; CAA70903.1; JOINED.
 DR EMBL: Y09765; CAA70904.1; -
 DR EMBL: Y07637; CAA68914.1; -
 DR EMBL: U92283; AAB94645.1; -
 DR EMBL: U92281; AAB94645.1; JOINED.
 DR EMBL: U92282; AAB94645.1; JOINED.
 DR Genew; HGNC:4085; GABRE.
 DR MIM: 300093; -
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.
 DR GO: GO:0004890; F: GABA-A receptor activity; TAS.
 DR GO: GO:000714; P: gamma-aminobutyric acid signaling pathway; TAS.
 DR GO: GO:0006832; P: small molecule transport; TAS.
 DR InterPro: IPR006029; Neu_channel memb.
 DR InterPro: IPR006202; Neu_channel memb.
 DR InterPro: IPR006201; Neu_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR TIGRFAMS: TIGR00860; LIC; 1.
 DR PROSITE: PS00236; NEUOTR_ION_CHANNEL; 1.
 DR Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KM Transmembrane. 1
 FT SIGNAL 23 506
 FT CHAIN 1
 FT DOMAIN 23 253
 FT TRANSSEM 254 274
 FT TRANSSEM 281 301
 FT TRANSSEM 344 364
 FT TRANSSEM 365 485
 FT TRANSSEM 486 506
 FT CARBOHYD 134 134
 FT CARBOHYD 252 252
 FT DISULFID 195 209
 FT CONFLICT 102 102
 FT CONFLICT 245 245
 FT CONFLICT 261 261
 FT CONFLICT 471 471
 FT CONFLICT 502 502
 SQ SEQUENCE 506 AA; 57971 MW; 013EE136EECFCC73 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 506;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPVNC 6
 DB 458 CWPVDC 463

RESULT 15
 ADEC_ARCFU STANDARD; PRT; 556 AA.
 AC 029999;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable adenine deaminase (EC 3.5.4.2) (Adenase) (Adenine aminase).
 GN AF0240.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; Pubmed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschess E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: BELONGS TO THE ADENINE DEAMINASE FAMILY.
 CC -1- CATALYTIC ACTIVITY: Adenine + H(2)O = hypoxanthine + NH(3).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF001089; AAB90993.1; -
 DR TIGR: H69279; H69279.
 DR PIR: AF0240; -
 DR InterPro: IPR006679; Adenine_deam.
 DR InterPro: IPR006680; Amidohydro_1.
 DR InterPro: IPR005847; Urease.
 DR Pfam: PF01979; Amidohydro_1; 1.
 DR Pfam: PF000518; Urease; 1.
 DR TIGRFAMS: TIGR01178; ade; 1.
 DR Hypothetical protein; Hydrolase; Complete proteome.
 SQ SEQUENCE 556 AA; 60764 MW; B3E3440A09F86A79 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 556;
 Best Local Similarity 66.7%; Pred. No. 63;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPVNC 6
 DB 137 CWPVSC 142

Search completed: September 4, 2003, 21:01:45
 Job time : 7.20588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 24.6176 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-60
Perfect score: 39
Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*
- 15: sp-virus:*
- 16: sp-bacteriaph:*
- 17: sp-archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	38	97.4	129	10	08H7Y2	08H7Y2 Oryza sativ
2	38	97.4	135	10	040228	040228 Liliium long
3	38	97.4	135	10	043538	043538 Liliium long
4	36	92.3	303	16	09HVR7	09HVR7 pseudomonas
5	35	89.7	288	5	09XVY5	09XVY5 toxocara ca
6	35	89.7	334	5	024403	024403 drosophila
7	35	89.7	334	5	09VAM8	09VAM8 drosophila
8	35	89.7	366	15	0997P7	0997P7 human immun
9	35	89.7	472	10	09MIP3	09MIP3 arabidopsis
10	35	89.7	505	16	09KSF2	09KSF2 vibrio chol
11	35	89.7	507	16	09HMF8	09HMF8 pseudomonas
12	35	89.7	507	16	08DAP0	08DAP0 vibrio vuln
13	35	89.7	613	5	09VGR8	09VGR8 drosophila
14	34	87.2	197	10	0944D8	0944D8 brassica ol
15	34	87.2	308	11	08CIT7	08CIT7 mus musculu
16	34	87.2	335	13	08JFU1	08JFU1 brachydanio

17	34	87.2	353	4	08IV68	08IV68 homo sapien
18	34	87.2	445	5	08T364	08T364 podocoryne
19	34	87.2	698	5	08IAO1	08IAO1 plasmodium
20	34	87.2	712	5	020323	020323 caenorhabdi
21	34	87.2	772	12	08JRS8	08JRS8 diachasma
22	34	87.2	1403	11	09J2D3	09J2D3 mus musculu
23	34	87.2	1898	11	08CHF2	08CHF2 mus musculu
24	34	87.2	2192	11	09J2U2	09J2U2 mus musculu
25	33	84.6	73	6	09N052	09N052 macaca fasc
26	33	84.6	86	9	094M96	094M96 lactococcus
27	33	84.6	99	4	09UJN6	09UJN6 homo sapien
28	33	84.6	105	4	060801	060801 homo sapien
29	33	84.6	106	4	0969Q0	0969Q0 homo sapien
30	33	84.6	106	13	09OVX4	09OVX4 brachydanio
31	33	84.6	126	6	09N0B8	09N0B8 macaca fasc
32	33	84.6	144	4	08NBL2	08NBL2 homo sapien
33	33	84.6	200	12	09IC10	09IC10 porcine rep
34	33	84.6	200	12	041198	041198 porcine rep
35	33	84.6	200	12	041196	041196 porcine rep
36	33	84.6	200	12	041193	041193 porcine rep
37	33	84.6	200	12	041192	041192 porcine rep
38	33	84.6	200	12	041199	041199 porcine rep
39	33	84.6	200	12	041195	041195 porcine rep
40	33	84.6	200	12	041188	041188 porcine rep
41	33	84.6	200	12	041197	041197 porcine rep
42	33	84.6	200	12	084936	084936 porcine rep
43	33	84.6	200	12	09J7J5	09J7J5 porcine rep
44	33	84.6	201	2	09ANB9	09ANB9 bradyrhizob
45	33	84.6	234	11	08KOT6	08KOT6 mus musculu

ALIGNMENTS

RESULT 1
ID 08H7Y2 PRELIMINARY; PRT; 129 AA.

AC 08H7Y2: 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OJ1607A12.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=3947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA Currie J., Collura K.
RT "Rice Genomic Sequence."
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105729; AAN06866.1;
KW Hypothetical protein.
SQ SEQUENCE 129 AA; 13261 MW; 775FE060768D4F82 CRC64;

Query Match 97.4%; Score 38; DB 10; Length 129;
Best Local Similarity 83.3%; Pred. No. 2.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 93 CVPNC 98

RESULT 2
ID 040228 PRELIMINARY; PRT; 135 AA.
AC 040228;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

```

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF (Fragment).
GN LIM5.
OS Lillium longiflorum (Trumpet Lily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lillium.
OX NCBI_TaxID=4690;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Flower buds;
RA Tabata S.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Flower buds;
RA MEDLINE=96051386; PubMed=7584025;
RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A.,
RA Tabata S.;
RT "Characterization of cDNAs induced in meiotic prophase in Lily
RT Microsporocytes."
RL DNA Res. 1:15-26(1994).
DR EMBL; D21811; BAA04835.1; -.
FT NON_TER
SQ SEQUENCE 135 AA; 14769 MW; D7E0AE91EPCD4E6B CRC64;

Query Match
Best Local Similarity 97.4%; Score 38; DB 10; Length 135;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 130 CVPNC 135

RESULT 3
ID 043538 PRELIMINARY; PRT; 135 AA.
AC 043538;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF (Fragment).
GN LIM13.
OS Lillium longiflorum (Trumpet Lily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lillium.
OX NCBI_TaxID=4690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hinomoto; TISSUE=Flower buds;
RA Tabata S.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hinomoto; TISSUE=Flower buds;
RA MEDLINE=96051386; PubMed=7584025;
RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A.,
RA Tabata S.;
RT "Characterization of cDNAs induced in meiotic prophase in Lily
RT Microsporocytes."
RL DNA Res. 1:15-26(1994).
DR EMBL; D21819; BAA04843.1; -.
FT NON_TER
SQ SEQUENCE 135 AA; 14858 MW; 283C7C532D91B818 CRC64;

Query Match
Best Local Similarity 97.4%; Score 38; DB 10; Length 135;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 130 CVPNC 135

```

```

RESULT 4
ID 09HVR7 PRELIMINARY; PRT; 303 AA.
AC 09HVR7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Probable permease of ABC transporter.
GN PA4504.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Gidder R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AEO04864; AMO7892.1; -.
DR InterPro: IPR000315; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 32367 MW; B9496E9EA55E82B3 CRC64;

Query Match
Best Local Similarity 92.3%; Score 36; DB 16; Length 303;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
DB 211 CVPNC 216

RESULT 5
ID 09XYV5 PRELIMINARY; PRT; 288 AA.
AC 09XYV5;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C-type lectin TC-cl-4.
OS Toxocara canis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OX NCBI_TaxID=6265;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21011334; PubMed=11128806;
RA Loukas A., Doedens A., Hintz M., Maizels R.M.;
RT "Identification of a new C-type lectin, TES-70, secreted by infective
RT larvae of Toxocara canis, which binds to host ligands."
RL Parasitology 121:545-554(2000).
DR EMBL; AF126830; AAD31000.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PSS0041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 288 AA; 30273 MW; 5C08AA9AA1311064 CRC64;

Query Match
Best Local Similarity 89.7%; Score 35; DB 5; Length 288;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 CVPVNC 6
Db      120 CAVPNC 125

RESULT 6
ID      024403      PRELIMINARY:      PRT:      334 AA.
AC      024403:
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Spem protein.
GN      MS798CA OR CG11719.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=OREGONR;
RA      Schaefer M., Boersch D., Huelster A., Schaefer U.;
RL      Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
DR      EMBL; X67704; CAA47941.1; -
DR      HSSP; P01180; INPO.
DR      FLYbase: FBgn0002865; Mst98CA.
DR      InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR      InterPro: IPR001368; TNFR_c6.
DR      PROSITE; PS00198; 4Fe4S_FERREDOXIN; 3.
DR      PROSITE; PS00652; TNFR_NGFR_1; 1.
SQ      SEQUENCE 334 AA; 34709 MW; 2C88361C34B9CCF CRC64;

Query Match
Best Local Similarity 89.7%; Score 35; DB 5; Length 334;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CVPVNC 6
Db      242 CAVPNC 247

RESULT 7
ID      09VAM8      PRELIMINARY:      PRT:      334 AA.
AC      09VAM8:
DT      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Mst98CA protein (GH09231P).
GN      Mst98CA OR CG11719.
OS      Drosophila melanogaster (Fruit fly)
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BERKELEY;
RA      MEDLINE=20196006; Pubmed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA      Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA      Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA      Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA      Dodson K., Doup L.E., Downes M., Dugan-Recha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA      Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reibert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA      "The genome sequence of Drosophila melanogaster."
RA      Science 287:2185-2195(2000).
RA      [2]
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Berkeley;
RA      Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA      Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA      Gonzalez M., Guarin H., Li P., Lao G., Miranda A., Mungall C.J.,
RA      Nunoo J., Pacle J., Paragas V., Park S., Phoumenavong S., Wan K.,
RA      Yu C., Lewis S.E., Rubin G.M., Park S., Phoumenavong S., Wan K.,
RA      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AE003764; AAF56779.1; -
DR      EMBL; AY060667; AAL28215.1; -
DR      HSSP; P01180; INPO.
DR      FLYbase: FBgn0002865; Mst98CA.
DR      InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR      InterPro: IPR001368; TNFR_c6.
DR      PROSITE; PS00198; 4Fe4S_FERREDOXIN; 3.
DR      PROSITE; PS00652; TNFR_NGFR_1; 1.
SQ      SEQUENCE 334 AA; 34725 MW; 2C876D1864BFC394 CRC64;

Query Match
Best Local Similarity 89.7%; Score 35; DB 5; Length 334;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CVPVNC 6
Db      242 CAVPNC 247

RESULT 8
ID      0997P7      PRELIMINARY:      PRT:      366 AA.
AC      0997P7:
DT      01-JUN-2001 (TREMBlrel. 17, Created)
DT      01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT      01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE      Envelope glycoprotein (Fragmint).
GN      Env.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Wang F., Kimura T., Nishihara K., Yoshimura K., Kotko A.,
RA      Matsushita S.;
RA      "Emergence of autologous neutralization-resistant variants from pre-
RA      RT existing quasiespecies during viral-rebound from patients under
RA      HAART."
RA      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```

DR EMBL: AB059310, BAB40894.1; --
 DR InterPro: IPR000777; gpi20.
 DR Pfam: PF00516; gpi20.1.
 DR Aids: Coat Protein; Glycoprotein.
 FT NON_TER 1
 SQ SEQUENCE 366 AA; 40154 MW; 287683B308FC8CB8 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 15; Length 366;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 CVPNC 6
 1 CVPNC 6

RESULT 9

AC 09M1P3 PRELIMINARY; PRT; 472 AA.
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-MAR-2003 (TREMblrel. 15, Last sequence update)
 DE Protein kinase-like protein.
 GN 118B22.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 OC euosids II; Brassicales; Brassicaceae; Arabidopsids.
 RX NCBI_TaxID:3702;
 RA Jordan N., Baugert S., Wiedemann R., Voss H., Unselid M., Mewes H.W.,
 RL Lemcke K., Baugert K.F.X., Queller F., Salanoubat M.;
 RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RT EU Arabidopsis sequencing project;
 CC EMBL: A1138652; BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; Pkinase.1.
 DR ProDom: PD000001; Pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SO SEQUENCE 472 AA; 52798 MW; 583042CFB904A3B3 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 10; Length 472;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 CVPNC 6
 1 CVPNC 6

DB 268 CVPNC 273
 RESULT 10
 ID 09KSF2 PRELIMINARY; PRT; 505 AA.
 AC 09KSF2
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE Putative hydrolase, class I, putative.
 GN V11304.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.

us-09-580-893d-60.ispt

OX NCBI_TaxID:666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-El Tor N1691 / Serotype O1;
 RX MEDLINE:20406833; PubMed:10952301;
 RA Heidelberg J.E., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Gill S.R., Haft D.H., Hickey E.K., Peterson J.D., Umhau L.A.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RT "DNA sequence of both chromosomes of the cholera pathogen vibrio
 choleraceae".
 RL Nature 406:477-483(2000).
 DR EMBL: AE004210; AAF94463.1;
 DR TIGR: VCI304;
 DR InterPro: IPR004646; TtdA_funa_fumb.
 DR TIGRFAMs: TIGR00722; TtdA_funa_fumb.
 DR TIGRFAMs: TIGR00723; TtdB_funa_fumb; 1.
 KW Complete proteome.
 SO SEQUENCE 505 AA; 54778 MW; 40C4A2C06449041F CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 16; Length 505;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 CVPNC 6
 1 CVPNC 6

DB 272 CVPNC 277
 RESULT 11
 ID 09HW68 PRELIMINARY; PRT; 507 AA.
 AC 09HW68
 DT 01-MAR-2001 (TREMblrel. 15, Created)
 DT 01-JUN-2002 (TREMblrel. 15, Last sequence update)
 DE Probable fumate hydratase.
 GN P4333.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 RX NCBI_TaxID:287;
 RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE:20437337; PubMed:10984043;
 RA Hickey M.J., Pham X.-O.T., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Brinkman F.S., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Brody L.L., Coulter S.N., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Kas A., Lartig K., Lim R.M.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Loy S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen".
 RL Nature 406:959-964(2000).
 DR EMBL: AE004848; AAG07721.1;
 DR InterPro: IPR004647; TtdA_funa_fumb.
 DR TIGRFAMs: TIGR00722; TtdA_funa_fumb.
 DR TIGRFAMs: TIGR00723; TtdB_funa_fumb; 1.
 KW Complete proteome.
 SO SEQUENCE 507 AA; 54763 MW; 49BCDF4A9F172907 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 16; Length 507;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 CVPNC 6
 1 CVPNC 6

DB 1 CVPNC 6
 1 CVPNC 6

Db 272 CIMPNC 277

RESULT 12

Q9DAE0 PRELIMINARY; PRT; 507 AA.

AC Q9DAE0; 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Fumarate hydratase.

CN V12266.

OS Vbrio vulnificus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrrio.

OX NCBI_TaxID=672;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CMCP6;

RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,

RA Chey H.E.;

RT "Complete genome sequence of *Vibrio vulnificus* CMCP6."

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AE016804; AAO10644.1; -

KW Complete proteome.

SO SEQUENCE 507 AA; 54848 MW; 210DF5328E5F2B7 CRC64;

Query Match 89.7%; Score 35; DB 16; Length 507;

Best Local Similarity 66.7%; Pred. No. 36;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPNC 6

Db 272 CIMPNC 277

RESULT 13

Q9VGR8 PRELIMINARY; PRT; 613 AA.

AC Q9VGR8; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE CG6689 protein.

CN CG6689.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,

RA Adili J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

DR EMBL: AE003690; AAF54607.1; -

DR FlyBase: FBgn0037877; CG6689.

DR InterPro: IPR006612; DUF_DM3.

DR InterPro: IPR007087; ZnF_C2H2.

DR InterPro: IPR006025; Zn_MTPepidase.

DR Pfam: PF00096; zf-C2H2_7.

DR ProDom: PD000003; ZnF_C2H2_1.

DR SMART: SM00355; ZnF_C2H2_7.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.

DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 7.

DR PROSITE: PS00142; ZINC_PROTEASE; 1.

KW Metal-binding; Zinc; Zinc-finger.

SO SEQUENCE 613 AA; 70431 MW; 8C9ADD2899C108B4 CRC64;

Query Match 89.7%; Score 35; DB 5; Length 613;

Best Local Similarity 83.3%; Pred. No. 43;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVPNC 6

Db 3 CAVPNC 8

RESULT 14

Q944D8 PRELIMINARY; PRT; 197 AA.

AC Q944D8; 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Resistance disease protein.

CN TUR7.

OS Brassica oleracea (Cauliflower).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI_TaxID=3712;

RN [1]

RP SEQUENCE FROM N.A.

RA Cao B., Lei J., Song H.;

RT "Tur7-resistance disease gene in *Brassica oleracea*."

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF432595; AAL28107.1; -

SO SEQUENCE 197 AA; 22932 MW; AE90AA8E2C23FA54 CRC64;

Query Match 87.2%; Score 34; DB 10; Length 197;

Best Local Similarity 83.3%; Pred. No. 25;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVPNC 6

Db 26 CVPNC 31

RESULT 15

Q8C1T7

ID Q8C1T7

PRELIMINARY;

PRT; 308 AA.

AC Q8C1T7;

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Spinal cord;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK090368; BAC41186.1; -.

KW Hypothetical protein.

FT NON_TER 308

SQ SEQUENCE 308 AA; 34483 MW; AD58F3417BB513D CRC64;

Query Match 87.2%; Score 34; DB 11; Length 308;

Best Local Similarity 83.3%; Pred. No. 37;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6

Db 84 CVCPCNC 89

Search completed: September 4, 2003, 21:06:43
Job time : 27.7086 secs

PI Miltis TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 43; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAVVPNC 7
Db 1 CAVVPNC 7
IIIIII
RESULT 2
AAE16729 standard; peptide; 7 AA.
XX
AC AAE16729;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #9.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
XX
XX Ligamentum nuchae.
OS Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note="This residue forms a cyclic structure with Cys
FT at position 7 with copper as a chelating agent"
FT Modified-site 7
FT /note="This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent."
XX
XX WO200191700-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 30-MAY-2001; 2001WO-US17384.
XX
XX PR 30-MAY-2000; 2000US-0580110.
XX PR 30-MAY-2000; 2000US-0580156.
XX PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.
XX
XX (CONN-) CONNECTIVE TISSUE IMAGING ENGINEERING LLC.
XX
PI Miltis TF, Sandberg LB, Jimenez F;
XX
DR WPI; 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
PT substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived cyclic peptide.
XX
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 43; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAVVPNC 7
Db 1 CAVVPNC 7
IIIIII
RESULT 3
AAE16715 standard; peptide; 7 AA.
XX
AC AAE16715;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPQ derived peptide #9.
XX
KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW hypotensive; chronic obstructive pulmonary disease; antarteriosclerotic;
KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
XX
XX Ligamentum nuchae.
OS Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..7
FT
XX
XX WO200191700-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 30-MAY-2001; 2001WO-US17384.
XX
XX PR 30-MAY-2000; 2000US-0580110.
XX PR 30-MAY-2000; 2000US-0580156.
XX PR 30-MAY-2000; 2000US-0580893.
XX PR 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING L.L.C.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI MPI; 2002-106259/14.
 DR Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX Claim 24; Page 20; 53pp; English.

CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.,
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.

SQ Sequence 7 AA:
 Query Match 86.0%; Score 37; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPC 7
 DB 1 CAVVPC 7

RESULT 4
 AAEL6718
 ID AAEL6718 standard; peptide; 7 AA.

AC AAEL6718;
 DT 09-APR-2002 (first entry)
 XX

DE Ligamentum nuchae hydrolysed elastin peptide, VVPQ derived peptide #12.
 XX
 KM Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KM vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KM hypotensive; chronic obstructive pulmonary disease; antileukosclerotic;
 KM restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KM cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX

EH Key 1 Location/Qualifiers
 FT Modified-site 1 /note="This residue forms a cyclic structure with Cys
 FT at position 7 with copper as a chelating agent."
 FT Modified-site 7 /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent."
 FT
 FT
 PN WO200191700-A2.
 XX
 XX 06-DEC-2001.
 PD
 XX 30-MAY-2001; 2001WO-US17384.
 PF
 XX

PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING L.L.C.
 XX Mlts TF, Sandberg LB, Jimenez F;
 PI MPI; 2002-106259/14.
 DR Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX Claim 24; Page 20; 53pp; English.

CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.,
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.

SQ Sequence 7 AA:
 Query Match 86.0%; Score 37; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPC 7
 DB 1 CAVVPC 7

RESULT 5
 AAU48588
 ID AAU48588 standard; Protein; 124 AA.

AC AAU48588;
 DT 27-FEB-2002 (first entry)
 XX

DE Propionibacterium acnes immunogenic protein #9484.
 XX
 KM SAPHO syndrome; synovitis; acne; pustulosis; hypertostis; osteomyelitis;
 KM uvelitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 OS
 PN WO200181581-A2.
 XX
 XX 01-NOV-2001.
 PD
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR MPI; 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PS
 PS Example 1; SEQ ID No 9783; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 124 AA;
 Query Match 83.7%; Score 36; DB 22; Length 124;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CAVVPCNC 7
 Db 87 CAVVPCNC 93
 RESULT 6
 ABP42962
 ID ABP42962 standard; Protein: 155 AA.
 AC
 AC ABP42962;
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HPV067, SEQ ID NO:4094.
 DE
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 XX
 PN WO200200677-A1.
 PD 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 PF
 XX 07-JUN-2000; 2000US-209467P.
 PR

XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 PI Blise CE, Rosen CA;
 XX
 DR MPI; 2002-147878/19.
 DR N-PSDB; AB056039.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 PS
 PS Claim 11; SEQ ID No 4094; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP4228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis), systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 155 AA;
 Query Match 81.4%; Score 35; DB 23; Length 155;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAVVPCNC 7
 Db 101 CAVVPCNC 107
 RESULT 7
 AAEL6727
 ID AAEL6727 standard; peptide: 8 AA.
 AC
 AC AAEL6727;
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, WVPN derive peptide #7.
 DE
 XX Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotonic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.
 OS

OS Synthetic.
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 CC
 XX
 SQ Sequence 8 AA;
 QY
 Query Match 79.1%; Score 34; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 2 AVVPNC 7
 3 AVVPNC 8
 XX
 RESULT 8
 AAE16730
 ID AAE16730 standard; peptide; 8 AA.
 XX
 AC AAE16730;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 1

FT /note= "This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"
 FT Modified-site 8
 FT /note= "This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI; 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived cyclic peptide.
 CC
 XX
 SQ Sequence 8 AA;
 QY
 Query Match 79.1%; Score 34; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 2 AVVPNC 7
 3 AVVPNC 8
 XX
 RESULT 9
 AAB25741
 ID AAB25741 standard; Protein; 67 AA.
 XX
 AC AAB25741;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 16 SEQ ID NO:130.
 XX
 KW Human; secreted protein; immunosuppressive; immunostimulant; nocotropic;
 KW antiinflammatory; cardiant; vulnerary; antitumor; anticonvulsant;
 KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
 KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
 KW gene therapy; vaccine; chemotaxis modulator; angiogenesis modulator;
 KW cancer; immune system disorder; hyperproliferative disorder; infection;
 KW cardiovascular disorder; neurological disease; wound healing.
 XX

XX	17-JAN-2001	2001MO-US0135
PR	31-JAN-2000	2000US-0179065
PR	04-FEB-2000	2000US-0180628
PR	24-FEB-2000	2000US-0184564
PR	02-MAR-2000	2000US-0186150
PR	16-MAR-2000	2000US-0189874
PR	17-MAR-2000	2000US-0190076
PR	18-MAR-2000	2000US-0191823
PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214886
PR	30-JUN-2000	2000US-0215335
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216880
PR	11-JUL-2000	2000US-0217487
PR	14-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	22-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226281
PR	22-AUG-2000	2000US-0226868
PR	23-AUG-2000	2000US-0227189
PR	30-AUG-2000	2000US-0227009
PR	01-SEP-2000	2000US-0228292
PR	01-SEP-2000	2000US-0228294
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	05-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	12-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232357
PR	14-SEP-2000	2000US-0232358
PR	14-SEP-2000	2000US-0232359
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	21-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234423
PR	21-SEP-2000	2000US-0234424
PR	25-SEP-2000	2000US-0234997
PR	25-SEP-2000	2000US-0234998
PR	27-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235485
PR	27-SEP-2000	2000US-0235536
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369

PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239393.
 PR 20-OCT-2000; 2000US-0239393.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250161.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251866.
 PR 08-DEC-2000; 2000US-0251867.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0253678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPT: 2001-483426/52.
 DR N-PSDB; AAK64053.
 XX

PT Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX
 PS Claim 11; SEQ ID NO 18865; 3071tp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (1)
 CC amino acid sequences given in AAM82170 to AAM91921. (1) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (1)
 CC CC
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (1) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (1) by expressing inactive proteins or to
 CC supplement the patient's own production of (1). Additionally, (1)
 CC polynucleotides may be used to produce the secreted (1), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (1) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 SQ Sequence 78 AA:

Query Match 79.1%; Score 34; DB 22; Length 78;
 Best Local Similarity 57.1%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
 I : : I I I
 DB 18 CLIPNC 24

RESULT 11
 AAM91271
 ID AAM91271 standard; Protein; 79 AA.

XX AAM91271;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:18864.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.

OS Homo sapiens.

PN WO200157182-A2.

XX 09-AUG-2001.

PD 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR	26-JUL-2000	2000US-02209664
PR	14-AUG-2000	2000US-02245119
PR	14-AUG-2000	2000US-02245119
PR	14-AUG-2000	2000US-02252113
PR	14-AUG-2000	2000US-02252113
PR	14-AUG-2000	2000US-02252667
PR	14-AUG-2000	2000US-02252679
PR	14-AUG-2000	2000US-02252679
PR	14-AUG-2000	2000US-02252686
PR	14-AUG-2000	2000US-02252686
PR	14-AUG-2000	2000US-02254270
PR	14-AUG-2000	2000US-02254471
PR	14-AUG-2000	2000US-02257558
PR	14-AUG-2000	2000US-02257558
PR	14-AUG-2000	2000US-02257577
PR	18-AUG-2000	2000US-02262729
PR	22-AUG-2000	2000US-02266611
PR	22-AUG-2000	2000US-02266611
PR	22-AUG-2000	2000US-02266688
PR	22-AUG-2000	2000US-02267182
PR	23-AUG-2000	2000US-02270190
PR	30-AUG-2000	2000US-02286924
PR	01-SEP-2000	2000US-02292827
PR	01-SEP-2000	2000US-02292827
PR	01-SEP-2000	2000US-02293344
PR	01-SEP-2000	2000US-02293344
PR	05-SEP-2000	2000US-02293505
PR	05-SEP-2000	2000US-02293505
PR	05-SEP-2000	2000US-02295137
PR	06-SEP-2000	2000US-02294513
PR	06-SEP-2000	2000US-02304338
PR	08-SEP-2000	2000US-02312442
PR	08-SEP-2000	2000US-02312443
PR	08-SEP-2000	2000US-02312444
PR	08-SEP-2000	2000US-02314113
PR	08-SEP-2000	2000US-02314114
PR	08-SEP-2000	2000US-02320800
PR	12-SEP-2000	2000US-02320861
PR	12-SEP-2000	2000US-02319688
PR	14-SEP-2000	2000US-02323297
PR	14-SEP-2000	2000US-02323298
PR	14-SEP-2000	2000US-02323299
PR	14-SEP-2000	2000US-02323299
PR	14-SEP-2000	2000US-02324011
PR	14-SEP-2000	2000US-02324011
PR	14-SEP-2000	2000US-02330663
PR	14-SEP-2000	2000US-02330664
PR	21-SEP-2000	2000US-02330655
PR	21-SEP-2000	2000US-02334223
PR	21-SEP-2000	2000US-02342724
PR	25-SEP-2000	2000US-02343997
PR	25-SEP-2000	2000US-02343998
PR	26-SEP-2000	2000US-02345834
PR	27-SEP-2000	2000US-02356584
PR	27-SEP-2000	2000US-02356586
PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363657
PR	29-SEP-2000	2000US-02363659
PR	29-SEP-2000	2000US-02365198
PR	02-OCT-2000	2000US-02366070
PR	02-OCT-2000	2000US-02366072
PR	02-OCT-2000	2000US-02370337
PR	02-OCT-2000	2000US-02370338
PR	02-OCT-2000	2000US-02370340
PR	13-OCT-2000	2000US-02393535
PR	13-OCT-2000	2000US-02393537
PR	20-OCT-2000	2000US-02409660
PR	20-OCT-2000	2000US-02412121
PR	20-OCT-2000	2000US-02417851
PR	20-OCT-2000	2000US-02417856
PR	20-OCT-2000	2000US-02417877
PR	20-OCT-2000	2000US-02418088
PR	20-OCT-2000	2000US-02418099
PR	01-NOV-2000	2000US-02441826
PR	08-NOV-2000	2000US-02446474
PR	08-NOV-2000	2000US-02464575

PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246479.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0250191.
PR 05-DEC-2000; 2000US-0250193.
PR 05-DEC-2000; 2000US-0251088.
PR 05-DEC-2000; 2000US-0251098.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX DR N-PSDB; AAK64052.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

XX Claim 11: SEQ ID NO 18864; 3071pp + Sequence Listing; English.

XX
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM62170 to AMM91921. (I) have cytostatic
CC activity and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC

CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX

SO Sequence 79 AA;

Query Match 79.1%; Score 34; DB 22; Length 79;
 Best Local Similarity 57.1%; Pred. No. 60;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
 I : : : : :
 18 CLIPNC 24

RESULT 12
 ABB09232
 ID ABB09232 standard; Protein; 119 AA.

AC ABB09232;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9223.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSF-) HYSFO INC.

PI Drmanac RT, Liu C, Tang YT;

DR N-P-SDB; AAS73419.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX

PS Claim 20; SEQ ID NO 39591; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABB00010-ABB30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 119 AA;

Query Match 79.1%; Score 34; DB 22; Length 119;
 Best Local Similarity 57.1%; Pred. No. 88;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
 I : : : : :
 108 CKTIPNC 114

RESULT 13
 ABB89784
 ID ABB89784 standard; Protein; 156 AA.

AC ABB89784;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 2160.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KM vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.

OS Homo sapiens.

PN WO200190304-A2.

PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR N-P-SDB; ABL90193.

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX

PS Claim 11; SEQ ID NO 2160; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL9449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 156 AA;

Query Match 79.1%; Score 34; DB 23; Length 156;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPC 7

Db 123 CAVVPC 129

RESULT 14

AAB25742 standard; Protein: 235 AA.

AAB25742;

04-DEC-2000 (first entry)

Human secreted protein sequence encoded by gene 16 SEQ ID NO:131.

Human; secreted protein; immunosuppressive; immunostimulant; neutrotic;
antiinflammatory; cardiant; vulnerary; antituler; anticonvulsant;
antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
antiparasitic; thrombolytic; anticoagulant; antidiabetic; cytostatic;
gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
cancer; immune system disorder; hyperproliferative disorder; infection;
cardiovascular disorder; neurological disease; wound healing.

Homo sapiens.

W0200043495-A2.

27-JUL-2000.

18-JAN-2000; 2000WO-US00903.

19-JAN-1999; 99US-0116330.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;

W01: 2000-499225/44.

New isolated polynucleotide encoding a secreted protein useful for
preventing, treating or ameliorating a medical condition -
Disclosure; Page 48; 451pp: English.

The polynucleotide sequences given in AAB25742 to AAB25748 encodes the
human secreted proteins given in AAB25655 to AAB25755. Human secreted
proteins have activities based on the tissues and cells the genes are
expressed in. Examples of activities include: immunosuppressive;
immunostimulant; antiinflammatory; cardiant; vulnerary; antituler;
neutrotic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
antibacterial; antiparasitic; thrombolytic; anticoagulant;
antidiabetic; cytostatic; gene therapy and as vaccines,
polynucleotides can be used in gene therapy and as vaccines,
chemotaxis-modulators and angiogenesis-modulators. The human secreted
proteins and polynucleotides can be used for diagnosing (the
susceptibility to) a pathological condition by determining the presence
or absence of a mutation in the polynucleotide or determining the
presence or amount of expression of the protein. The polynucleotides and
proteins can also be used in the treatment and diagnosis of cancer,
diseases of the immune system, hyperproliferative disorders,
cardiovascular disorders and neurological disease. They can also be used
to promote wound healing and to fight infection. AAB25655 to AAB25664
represent sequences used in the exemplification of the present
invention.

SQ Sequence 235 AA;

Query Match 79.1%; Score 34; DB 21; Length 235;

Best Local Similarity 57.1%; Pred. No. 1.7e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVPC 7

Db 52 CVTIPNC 58

RESULT 15

AAE22089 standard; Protein: 316 AA.

AAE22089;

25-JUL-2002 (first entry)

Human novel ion exchanger protein (NHP) #2.

Human; novel ion exchanger protein; NHP; gene mapping; forensic biology;
restriction fragment length polymorphism; RFLP analysis; chromosome 14.

Homo sapiens.

W0200226980-A2.

04-APR-2002.

24-SEP-2001; 2001WO-US29828.

27-SEP-2000; 2000US-235745P.

(LEXI-) LEXICON GENETICS INC.

Fridde CJ, Gerhardt B;

W01: 2002-362496/39.

N-PSDB: AAD35145.

New isolated nucleic acid molecule for encoding a novel human ion
exchanger protein -

Claim 1; Page 37-38; 40pp: English.

The invention relates to novel human ion exchanger proteins (NHP) and
their corresponding nucleic acid sequences. The nucleic acid sequence of
the invention is used for encoding a novel human ion exchanger protein,
and is useful for the identification of a protein coding sequence and
mapping a unique gene to a particular chromosome. It may also be used
as additional DNA markers for restriction fragment length polymorphism
(RFLP) analysis, and in forensic biology. The present sequence is human
NHP protein. NHP gene is located on chromosome 14.

SQ Sequence 316 AA;

Query Match 79.1%; Score 34; DB 23; Length 316;

Best Local Similarity 57.1%; Pred. No. 2.2e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVPC 7

Db 132 CVTIPNC 138

Search completed: September 4, 2003, 21:00:34
Job time : 37.0294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 10.9118 seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893d-61

Perfect score: 43

Sequence: 1 CAVVPNC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR76:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	86.0	409	2 F90231	UDP-glucose 6-dehy
2	34	79.1	368	2 B70853	probable adhD prot
3	34	79.1	375	2 AB3618	alcohol dehydrogen
4	34	79.1	465	2 A56679	probable RNA-bind
5	34	79.1	644	2 T33132	hypothetical prote
6	34	79.1	784	2 J00317	hypothetical 82K p
7	33	76.7	127	2 S47895	SOH1 protein - yea
8	33	76.7	400	2 C70837	hypothetical prote
9	33	76.7	449	2 B71917	udp-n-acetylmuram
10	33	76.7	594	2 T23841	hypothetical prote
11	33	76.7	610	2 T18441	asparagine synthas
12	33	76.7	610	2 T23836	hypothetical prote
13	33	76.7	728	1 JH0579	hepatocyte growth
14	33	76.7	909	2 T06246	aspartate kinase (
15	33	76.7	916	2 T06242	aspartate kinase (
16	32	74.4	157	2 S41891	trypthophan 2-mono
17	32	74.4	174	2 T11844	NADH2 dehydrogenas
18	32	74.4	180	2 A45810	glycoprotein antiG
19	32	74.4	196	2 S73870	MG208 homolog H10
20	32	74.4	379	2 S42543	hypothetical prote
21	32	74.4	519	2 D71881	phosphotransacetyl
22	32	74.4	690	2 F70628	probable pta prote
23	32	74.4	755	1 DA6GWT	trypthophan 2-mono
24	32	74.4	755	1 QA6G4T	trypthophan 2-mono
25	32	74.4	755	2 A13328	trypthophan 2-mono
26	32	74.4	837	2 A34898	granulocyte colony
27	32	74.4	865	2 T41909	hypothetical prote
28	32	74.4	1014	2 T20361	hypothetical prote
29	31.5	73.3	1167	2 T06146	disease resistance

30	31	72.1	113	2 T24164	hypothetical prote
31	31	72.1	158	2 T26692	hypothetical prote
32	31	72.1	166	2 T28809	hypothetical prote
33	31	72.1	174	2 T14027	NADH2 dehydrogenas
34	31	72.1	174	2 C59154	NADH2 dehydrogenas
35	31	72.1	246	1 S29279	acetyl-CoA re
36	31	72.1	274	2 D72044	hypothetical prote
37	31	72.1	274	2 A86581	hypothetical prote
38	31	72.1	334	2 AG2172	hypothetical prote
39	31	72.1	366	2 T39804	hypothetical prote
40	31	72.1	436	2 T28066	hypothetical prote
41	31	72.1	449	2 S76839	hypothetical prote
42	31	72.1	461	2 AB2185	hypothetical prote
43	31	72.1	469	2 T46929	hypothetical prote
44	31	72.1	469	2 T46930	hypothetical prote
45	31	72.1	511	2 JC7692	oligodendrocyte tr

ALIGNMENTS

RESULT 1
F90231
UDP-glucose 6-dehydrogenase (ugd) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: F90231
R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
Submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90231
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <KUR>
A:Cross-references: GB:AE006641; NID:g13813993; PIDN:AAK41109.1; GSPDB:GN00155
C:Genetics:
A:Gene: ugd
C:Superfamily: GDPmannose dehydrogenase

Query Match 86.0%; Score 37; DB 2; Length 409;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPNC 7
Db 213 CEVIPNC 219

RESULT 2
B70853
probable adhD protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70853
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MIMD:98295987; PMID:9634230
A:Accession: B70853
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-368 <COL>
A:Cross-references: GB:AL021309; GB:AL123456; NID:g3261510; PIDN:CAA16144.1; PID:g280
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: adhD
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

F:25-359/Domain: long-chain alcohol dehydrogenase homology <LADH>

Query Match 79.1%; Score 34; DB 2; Length 368;
Best Local Similarity 57.1%; Pred. No. 54;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
| : |||
Db 85 CSFPMC 91

RESULT 3

AB3618

alcohol dehydrogenase (EC 1.1.1.1) [Imported] - *Brucella melitensis* (strain 16M)

C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AB3618
R:Delvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova, I.; Mazur, M.; Goldsman, E.; Salkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leclercq, A.; Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3618

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <ROR>
A:Cross-references: GB:AE008918; PIDN:AAL54109.1; PID:q17985069; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10867
A:Map position: 11
C:Superfamily: alcohol dehydrogenase: long-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase

Query Match

79.1%; Score 34; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
| : |||
Db 93 CVFPMC 99

RESULT 4

A56679

probable RNA-binding protein - house mosquito

C:Species: *Culex pipiens* (house mosquito)

C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jul-2000

C:Accession: A56679
R:Agarwal, M.; Bensaadi, N.; Salgado, J.C.; Campbell, K.; Mouches, C.
Insect Biochem. Mol. Biol. 23, 621-629, 1993

A:Title: Characterization and genetic organization of full-length copies of a LINE retrovirus
A:Reference number: A56679; MUID:93357789; PMID:8394766
A:Accession: A56679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <AGA>

A:Cross-references: GB:M01082; NID:9403079; PIDN:AAA28290.1; PID:9403080

Query Match 79.1%; Score 34; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
| : |||
Db 267 CAVPMC 273

RESULT 5

T33132

hypothetical protein F41A4.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33132
R:Bradshaw, H.; McPherson, C.
Submitted to the EMBL Data Library, May 1998
A:Description: The sequence of *C. elegans* cosmid F41A4.
A:Reference number: 221287
A:Accession: T33132

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-644

A:Cross-references: EMBL:AF067610; PIDN:AC17538.1; GSPDB:GN00022; CESP:F41A4.1
A:Experimental source: strain Bristol N2; clone F41A4
C:Genetics:
A:Gene: CESP:F41A4.1
A:Map position: 4
A:Introns: 57/3; 106/1; 147/2; 195/1; 244/2; 288/1; 308/2; 345/3; 374/3; 458/3

Query Match 79.1%; Score 34; DB 2; Length 644;
Best Local Similarity 71.4%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
| : |||
Db 55 CSSVPMC 61

RESULT 6

J00317

hypothetical 82K protein - *Xanthomonas campestris* pv. *vesicatoria*

C:Species: *Xanthomonas campestris* pv. *vesicatoria*
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993

C:Accession: J00317
R:Bonas, U.; Stall, R.E.; Staskiewicz, B.
Mol. Gen. Genet. 218, 127-136, 1989

A:Title: Genetic and structural characterization of the avirulence gene *avrBs3* from *X. campestris* pv. *vesicatoria*
A:Reference number: J00317; MUID:89384426; PMID:2550761
A:Accession: J00317
A:Molecule type: DNA
A:Residues: 1-784 <BON>

A:Comment: *X. campestris* pv. *vesicatoria* is the causal agent of bacterial leaf spot d
C:Genetics:
A:Start codon: GTG

Query Match 79.1%; Score 34; DB 2; Length 784;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
| : |||
Db 13 CULPNC 19

RESULT 7

S47895

SOH1 protein - Yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein G2864; protein YG1127c

C:Species: *Saccharomyces cerevisiae*
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 21-Jul-2000

C:Accession: S47895; S64138; S63786
R:Fan, H.; Klein, H.L.
Genetics 137, 945-956, 1994

A:Title: Characterization of mutations that suppress the temperature-sensitive growth of *S. cerevisiae*
A:Reference number: S47895; MUID:95073608; PMID:7982575
A:Accession: S47895
A:Molecule type: DNA
A:Residues: 1-127 <FAN>

A:Cross-references: EMBL:L131921; NID:9531011; PIDN:AAA35066.1; PID:9531012
R:Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Bejonte, E.; Tizon, B.; Cadahia, J.L.
Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64134
A:Accession: S64134
A:Molecule type: DNA
A:Residues: 1-127 <CER>

A:Cross-references: EMBL:272649; NID:91322690; PID:91322691; MIPS:YGL127C

A:Experimental source: strain S286C
 R:Rodriguez-Belmonte, E.; Rodriguez-Torres, A.M.; Tizon, B.; Cadahia, J.L.; Gonzalez-Sis
 Yeast 12, 145-148, 1996
 A:Title: Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII reveals a r
 A:Reference number: S63783; MUID:96287651; PMID:8666378
 A:Accession: S63786
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-127 <ROD>
 A:Cross-references: EMBL:X87252; NID:g1246906; PIDN:CA60704.1; PID:g1246910
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
 C:Genetics:
 A:Gene: SGD:SOH1
 A:Cross-references: SGD:S0003095; MIPS:YGL127C
 A:Map position: 7L
 C:Superfamily: Saccharomyces SOH1 protein
 C:Keywords: transmembrane protein
 F:78-94/Domain: transmembrane #status predicted <TM>

Query Match 76.7%; Score 33; DB 2; Length 127;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 11:111
 Db 72 CAVVNC 78

RESULT 8

C70837
 hypothetical protein RV0293c - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: C70837
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
 ; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: C70837
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-400 <COL>
 A:Cross-references: GB:AL021930; GB:AL123456; NID:g3261524; PIDN:CA17368.1; PID:e125247
 C:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: RV0293c

Query Match 76.7%; Score 33; DB 2; Length 400;
 Best Local Similarity 83.3%; Pred. No. 89;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AVVNC 7
 11:111
 Db 283 AVVNC 288

RESULT 9

B71917
 udp-n-acetylmuramate-alanine ligase - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
 C:Accession: B71917
 R:Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923662
 A:Accession: B71917
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-449 <ARN>
 A:Cross-references: GB:AF001489; GB:AF001439; NID:g4155102; PIDN:AA006138.1; PID:g4415
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: murC
 C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 76.7%; Score 33; DB 2; Length 449;
 Best Local Similarity 83.3%; Pred. No. 98;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAVVNC 6
 11:111
 Db 180 CAVVNC 185

RESULT 10

T23841
 hypothetical protein M88.6b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T23841
 R:Sulston, J.
 submitted to the EMBL Data Library, June 1994
 A:Reference number: Z19806
 A:Accession: T23841
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-594 <WIL>
 A:Cross-references: EMBL:Z34802; PIDN:CAB54282.1; GSPDB:GN00021; CESP:M88.6b
 A:Experimental source: clone M88
 C:Genetics:
 A:Gene: CESP:M88.6b
 A:Map position: 3
 A:Insertions: 40/3; 60/2; 101/3; 216/2; 290/1; 361/1; 411/3; 455/2; 544/3

Query Match 76.7%; Score 33; DB 2; Length 594;
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 11:111
 Db 488 CLAIPC 494

RESULT 11

T18441
 asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [similarity] - malaria paras
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-Jul-2002
 C:Accession: T18441
 R:Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z18935
 A:Accession: T18441
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-610 <LAW>
 A:Cross-references: EMBL:Z98547; PIDN:CAB1114.1
 C:Genetics:
 A:Map position: 3
 A:Insertions: 27/2
 A:Note: C0395W
 C:Superfamily: asparagine synthase (glutamine-hydrolyzing)
 C:Keywords: asparagine biosynthesis; ligase
 F:2-610/Product: asparagine synthase (glutamine-hydrolyzing) #status predicted <DUM>
 F:2/Active site: Cys #status predicted

Query Match 76.7%; Score 33; DB 2; Length 610;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAVYPN 6
DB 107 CAVYPN 112

RESULT 12
R:23836
hypotheoretical protein M88.6a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23836
R:Stulston, J.
submitted to the EMBL Data Library, June 1994
A:Reference number: Z19806
A:Accession: T23836
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-610 <MW>
A:Cross-references: EMBL:Z34802; PIDN:CAA84337.1; GSPDB:GND0021; CESP:M88.6a
A:Experimental source: clone M88
C:Genetics:
A:Gene: CESP:M88.6a
A:Map position: 3
A:Introns: 40/3; 60/2; 101/3; 216/2; 290/1; 427/3; 471/1; 560/3
Query Match 76.7%; Score 33; DB 2; Length 610;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVYPNC 7
DB 504 CLAYPNC 510

RESULT 13
JH0579
hepatocyte growth factor precursor [validated] - human
N:Alternate names: hepatocytein A; scatter factor
C:Species: Homo sapiens (man)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C:Accession: JH0579; JH0333; A41140; B36677; A33512; A39006; PH0114; A37796; S06
R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A:Title: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0579; MUID:91340155; PMID:1831432
A:Accession: JH0579
A:Molecule type: DNA
A:Residues: 1-728 <SEK>
A:Cross-references: DDBJ:D90318
R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
submitted to JIPID, March 1991
A:Description: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0333
A:Accession: JH0333
A:Molecule type: DNA
A:Residues: 1-481, 'RT', 484-728 <SE2>
R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor.
A:Reference number: A41140; MUID:91334393; PMID:1831266
A:Accession: A41140
A:Molecule type: mRNA
A:Residues: 1-728 <WEI>
A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936
R:Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor.
A:Reference number: A36677; MUID:91025062; PMID:2145836
A:Accession: B36677
A:Molecule type: mRNA
A:Residues: 1-728 <SE3>
A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032

A:Accession: A36677
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <SE4>
A:Cross-references: EMBL:X16323
A:Experimental source: leukocyte
R:Miyaizawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
A:Reference number: A33512; MUID:89392017; PMID:2528952
A:Accession: A33512
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-728 <MTY>
A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306846
R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.;
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepat
A:Reference number: A39006; MUID:91110540; PMID:1824873
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <RUB>
A:Cross-references: GB:M53579
A:Experimental source: embryonic lung
R:Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayam
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native a
A:Reference number: PH0114; MUID:91207365; PMID:1826837
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43; 53-58 <YOS>
A:Experimental source: plasma
R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of
A:Reference number: A37796; MUID:91035621; PMID:2146276
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91; 329-344; 356-363, 'XX', 366-370; 425-434; 442-447, 'X', 449-450; 543-546, 'X'
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Ta
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: S06794; MUID:90066676; PMID:2531289
A:Accession: S06794
A:Molecule type: mRNA
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-38
A:Cross-references: EMBL:X16333; NID:g32081; PIDN:CAA34387.1; PID:g32082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
R:Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Bir
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth f
A:Reference number: I59214; MUID:93087571; PMID:1280830
A:Accession: I59214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-288, 'ET', <HAR>
A:Cross-references: GB:I02931; NID:g184033; PIDN:AAA52649.1; PID:g184034
R:Miyaizawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth facto
A:Reference number: S15443; MUID:91200041; PMID:1826653
A:Accession: S15443
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288, 'ET', <MTY2>
A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: clon
A:Reference number: I52253; MUID:92062058; PMID:1835383
A:Accession: I52253
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 161-166 <SH1>
 A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997
 C:Genetics:
 A:Gene: GDB:HGR
 A:Cross-references: GDB:127524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A>Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringie homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringie; pyr
 F:1.31/Domain: signal sequence #status predicted <SIG>
 F:32-494/495-728/Product: hepatocyte growth factor #status experimental <AACP>
 F:32-494/Domain: alpha chain #status experimental <AACP>
 F:128-206/Domain: kringie homology <KR1>
 F:211-288/Domain: kringie homology <KR2>
 F:305-383/Domain: kringie homology <KR3>
 F:391-469/Domain: kringie homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCB>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F:294,402,566,653/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:487-604/Disulfide bonds: #status predicted

Query Match 76.7%; Score 33; DB 1; Length 728;
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 Db 377 CSQIPNC 383

RESULT 14
 T06246
 aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - soybean
 N:Contains: aspartate kinase (EC 2.7.2.4); homoserine dehydrogenase (EC 1.1.1.3)
 C:Species: Glycine max (soybean)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
 C:Accession: T06246
 R:Gebhardt, J.S.; Weisemann, J.M.; Matthews, B.F.
 submitted to the EMBL Data Library, February 1998
 A:Description: Genes encoding the bifunctional aspartokinase-homoserine dehydrogenase fr
 A:Reference number: Z15563
 A:Accession: T06246
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-909 <GERB>
 A:Cross-references: EMBL:AF049708; NID:g2970554; PIDN:AAC05983.1; PID:g2970556
 A:Experimental source: Cultivar Century
 C:Genetics:
 A:Gene: AK-HSDH
 A:Introns: 47/2; 69/1; 194/1; 223/3; 325/1; 342/3; 414/2; 443/3; 481/3; 514/3; 579/3; 60
 C:Function: <AK>
 A:Description: catalyzes phosphorylation of aspartate
 C:Function: <HSD>
 A:Description: catalyzes reduction of aspartate beta-semialdehyde into homoserine
 C:Superfamily: thra bifunctional enzyme; aspartate kinase homology; homoserine dehydr
 C:Keywords: multifunctional enzyme; oxidoreductase; phosphotransferase
 F:87-552/Domain: aspartate kinase homology <DK1>
 F:555-813/Domain: homoserine dehydrogenase homology <HSD>

Query Match 76.7%; Score 33; DB 2; Length 909;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

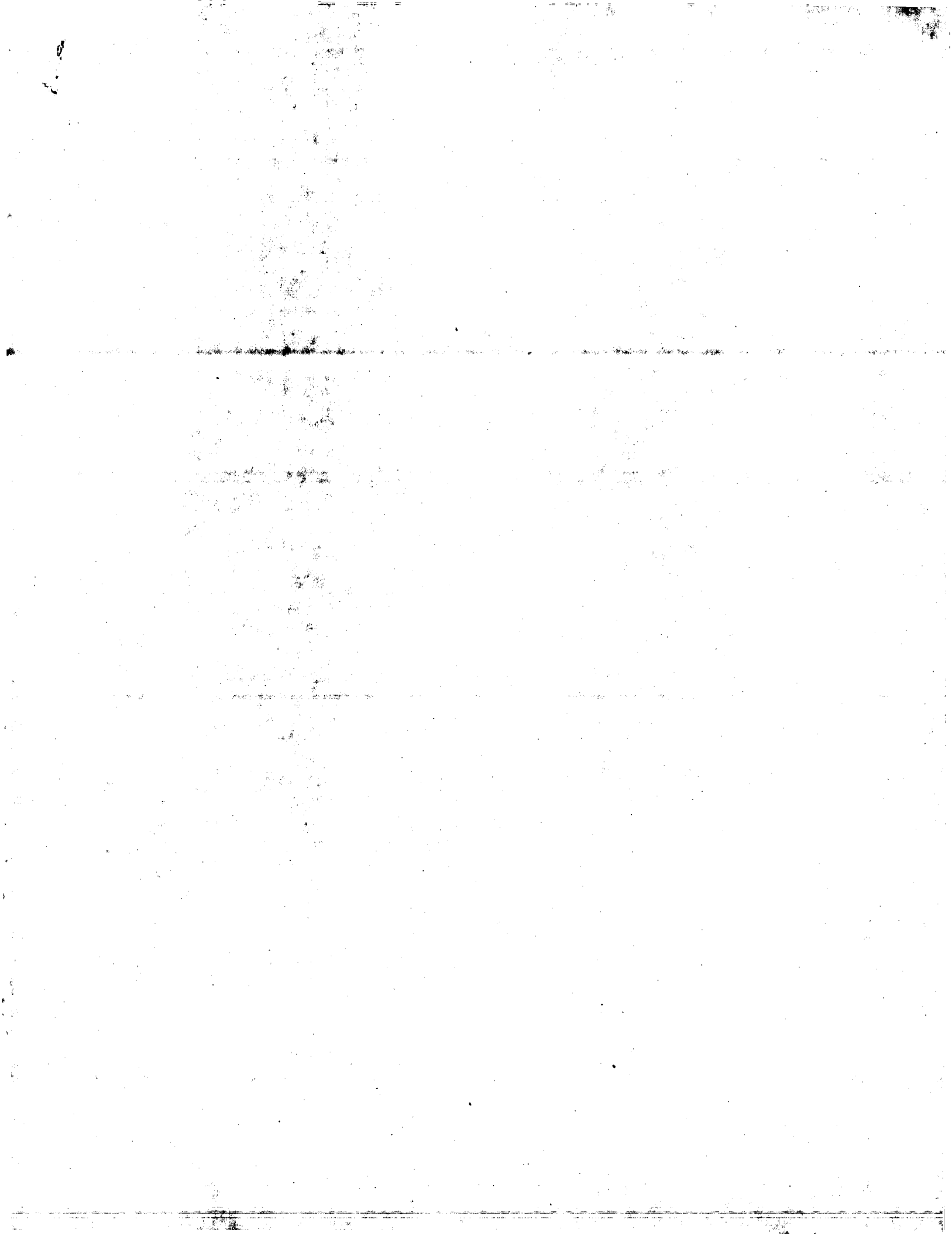
OY 2 AVVPNC 7
 Db 483 AVIPNC 488

RESULT 15
 T06242
 aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - soy
 N:Contains: aspartate kinase (EC 2.7.2.4); homoserine dehydrogenase (EC 1.1.1.3)
 C:Species: Chloroplast Glycine max (soybean)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
 C:Accession: T06242
 R:Gebhardt, J.S.; Weisemann, J.M.; Matthews, B.F.
 submitted to the EMBL Data Library, February 1998
 A:Description: Genes encoding the bifunctional aspartokinase-homoserine dehydrogenase
 A:Reference number: Z15563
 A:Accession: T06242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-916 <GERB>
 A:Cross-references: EMBL:AF049706; NID:g2970446; PIDN:AAC05981.1; PID:g2970447
 A:Experimental source: Cultivar Century
 C:Genetics:
 A:Gene: AK-HSDH
 A:Genome: chloroplast
 C:Function: <AK>
 A:Description: catalyzes phosphorylation of aspartate
 C:Function: <HSD>
 A:Description: catalyzes reduction of aspartate beta-semialdehyde into homoserine
 C:Superfamily: thra bifunctional enzyme; aspartate kinase homology; homoserine dehydr
 C:Keywords: chloroplast; multifunctional enzyme; oxidoreductase; phosphotransferase
 F:87-552/Domain: aspartate kinase homology <DK1>
 F:555-813/Domain: homoserine dehydrogenase homology <HSD>

Query Match 76.7%; Score 33; DB 2; Length 916;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AVVPNC 7
 Db 483 AVIPNC 488

Search completed: September 4, 2003, 21:10:51
 Job time : 11.9118 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 : Search time 6.07353 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-61
Perfect score: 43
Sequence: 1 CAVVPMC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	34	79.1	YAV2_XANCV	P14728 xanthomonas
2	33	76.7	SOH1_YEAST	P38633 saccharomyc
3	33	76.7	FSL3_MOUSE	Q99677 mus musculi
4	33	76.7	MURC_HELPJ	Q92112 helicobacte
5	33	76.7	HGF_HUMAN	P14210 homo sapien
6	32	74.4	NU6M_HYLLA	Q95710 hylobates 1
7	32	74.4	NU6M_PONPA	P92700 pongo pygma
8	32	74.4	Y208_MGCPN	P75486 mycoplasma
9	32	74.4	HMB1_S0YBN	P46608 glycolase max
10	32	74.4	PTA_HELPJ	Q92k44 helicobacte
11	32	74.4	PTA_MYCTU	P96254 mycobacteri
12	32	74.4	TR2M_AGR4	P04029 agrobacteri
13	32	74.4	TR2M_AGR4	P25017 agrobacteri
14	32	74.4	GCSR_MOUSE	P40223 mus musculi
15	32	74.4	VU5_HSV7J	P32522 human herpe
16	31	72.1	NU6M_GORGO	Q34573 gorilla gor
17	31	72.1	PHBB_CHRVI	P45375 chromatiu
18	31	72.1	Y753_SYNY3	P46535 synochocyst
19	31	72.1	Y920_TREPA	Q83880 tireponema p
20	31	72.1	FAT2_DROME	Q9W471 dtrosophila
21	30	69.8	A63F_DROME	Q46200 dtrosophila
22	30	69.8	TFE1_RAT	Q63467 rattus norv
23	30	69.8	PTA_CLOFM	Q52593 clostridium
24	30	69.8	ARPE_CHICK	Q9dee9 gallus gall
25	30	69.8	ARPE_HUMAN	Q9dz1 homo sapien
26	30	69.8	ARPE_MOUSE	Q9d864 mus musculi
27	30	69.8	ACTU_DROME	P45850 dtrosophila
28	30	69.8	NIFS_AZOB	P70727 azospirillum
29	30	69.8	ARGJ_STAAW	Q8nym7 s arginine
30	30	69.8	LY41_THEMA	Q9wz22 thermotoga
31	30	69.8	SYE_AERPE	Q9y9h1 aeropyrum p
32	30	69.8	DL_DROME	P10041 dtrosophila
33	30	69.8	ARPE_YEAST	Q12386 saccharomyc

34	30	69.8	1146	1	AGE1_CAEEL	Q94125 caenorhabdi
35	29	67.4	72	1	HPM1_ECTVA	P38941 ectothiorho
36	29	67.4	122	1	YPM6_YEAST	P43624 saccharomyc
37	29	67.4	170	1	YME1_THIFE	P22902 thlobacilli
38	29	67.4	175	1	PEP4_DROME	P34194 dtrosophila
39	29	67.4	208	1	HIS7_TRIHA	P34041 trichoderma
40	29	67.4	220	1	YHCW_BACSU	P54607 bacillus su
41	29	67.4	227	1	HIS7_MAGCR	Q42621 magnaporthe
42	29	67.4	271	1	RCE1_SCHPO	Q94448 schizosacch
43	29	67.4	293	1	RCE6_DROME	Q9vd25 dtrosophila
44	29	67.4	302	1	WR40_ARATH	Q9sah7 arabidopsis
45	29	67.4	330	1	UPAR_BOVIN	Q05588 bos taurus

ALIGNMENTS

RESULT 1
YAV2_XANCV
ID YAV2_XANCV STANDARD: PRT: 784 AA.
AC P14728:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 82 kDa avirulence protein in avrBs3 region.
OS Xanthomonas campestris (pv. vesicatoria).
OG Plasmid pXV11.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=341;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate 71-21;
RX MEDLINE=69384426; PubMed=2550761.
RA Bonas U., Stall R.E., Staskavicz B.;
RT "Genetic and structural characterization of the avirulence gene
avrb3 from Xanthomonas campestris pv. vesicatoria.";
RL Mol. Gen. Genet. 218:127-136(1989).
CC -1- DISEASE: THE TERM AVIRULENT DESCRIBES A POTENTIALLY VIRULENT
PATHOGEN WHICH IS UNABLE TO INDUCE DISEASE SYMPTOMS IN SPECIFIC
CULTIVARS OF PEPPER.
CC -1- MISCELLANEOUS: THIS IS ONE OF THE HYPOTHETICAL PROTEINS CODED BY
THE OPEN READING FRAMES WITHIN THE REGION REQUIRED FOR AVRBS3
ACTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X16130; -; NOT_ANNOTATED_CDS.
DR PIR: J00317; J00317.
DR InterPro: IPR005042; Avirulence.
DR Pfam: PF03377; Avirulence; 1.
KW Hypothetical protein; Plasmid; Virulence.
SO SEQUENCE 784 AA; 82074 MW; 3B2624BA1193744 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 784;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPMC 7
DB 13 CLIPNC 19

RESULT 2
SOH1_YEAST
ID SOH1_YEAST STANDARD: PRT: 127 AA.
AC P38633;

```

DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SOH1 protein.
OS SOH1 OR YGL127C OR G2864.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M303;
RX MEDLINE=95073608; PubMed=7982575;
RA Fan H.-Y., Klein H.L.;
RT "Characterization of mutations that suppress the temperature-sensitive
RT growth of the hpr1 delta mutant of Saccharomyces cerevisiae.";
RL Genetics 137:945-956(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96287651; PubMed=8686378;
RA Rodriguez-Belmonte E., Rodriguez Torres A.M., Tizon B., Cadahia J.L.,
RA Gonzalez-Siso I., Ramil E., Becerra M., Gonzalez-Dominguez M.,
RA Cerdan E.;
RT "Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII
RT reveals a novel member of the DnaJ family.";
RL Yeast 12:145-148(1996).
CC -1- FUNCTION: SUPPRESSES THE TEMPERATURE SENSITIVE GROWTH OF HPR1
CC MUTANT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the mediator SOH1 subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L31921; AAA35066.1; -;
DR EMBL: X87252; CA60704.1; -;
DR EMBL: Z72649; CA96836.1; -;
DR PIR: S47895; S47895.
DR TRANSFAC: T03615; -.
DR SGD: S0003095; SOH1.
DR GO: GO:0006281; P:DNA repair; IPL.
DR GO: GO:0006366; P:transcription from Pol II promoter; IGI.
DR Transcription regulation; Nuclear protein.
KW SEQUENCE 127 AA; 14741 MW; A2FE3B62ADC14011 CRC64;
SO
Query Match 76.7%; Score 33; DB 1; Length 127;
Best Local Similarity 71.4%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CAVVNC 7
DB 72 CIVVNC 78

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20576231; PubMed=11010968;
RA Tsuchida K., Arai K.Y., Kuramoto Y., Yamakawa N., Hasegawa Y.,
RA Sugino H.;
RT "Identification and characterization of a novel follistatin-like
RT protein as a binding protein for the TGF-beta family.";
RL J. Biol. Chem. 275:40788-40796(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Uchiyama Y., Ohnawa Y., Kametaka S.;
RT "Mouse follistatin-related protein FLRG (mouse PCr35, PC12 cell
RT derived trophic factor 35)";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21352723; PubMed=11459787;
RA Tortorello D.V., Sidis Y., Holtzman D.A., Holmes W.E.,
RA Schneyer A.L.;
RT "Human follistatin-related protein: a structural homologue of
RT follistatin with nuclear localization.";
RL Endocrinology 142:3426-3434(2001).
CC -1- FUNCTION: Binding protein for activin and BMP-2.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: Abundantly expressed in heart, lung, kidney,
CC and testis.
CC -1- SIMILARITY: Contains 2 follistatin-like domains.
CC -1- SIMILARITY: Contains 2 Kazal-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF276238; AAC47666.1; -;
DR EMBL: AB024429; BAB32663.1; -;
DR EMBL: AF310616; AAG53666.1; -;
DR HSP: P37109; 1PCE.
DR MGD: MGI:1890391; Fsl13.
DR InterPro: IPR003645; FOLN.
DR InterPro: IPR002350; kazal.
DR Pfam: PF00050; kazal; 2.
DR SMART: SM00274; FOLN; 2.
DR SMART: SM00280; KAZAL; 2.
DR PROSITE: PS00282; KAZAL; FALSE_NEG.
KW Glycoprotein; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 256
FT DOMAIN 96 118
FT DOMAIN 119 165
FT DOMAIN 168 191
FT DOMAIN 193 241
FT DOMAIN 241 256
FT DISULFID 119 151
FT DISULFID 123 144
FT DISULFID 133 165
FT DISULFID 193 227
FT DISULFID 198 220
FT DISULFID 209 241
FT CARBOHYD 71 71
FT CARBOHYD 213 213
FT SEQUENCE 256 AA; 27271 MW; 5854A4B8C10C2EE9 CRC64;
SO
Query Match 76.7%; Score 33; DB 1; Length 256;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CAVVNC 7
DB 117 CECVNC 123

```

```

RESULT 4
MURC_HELPJ
ID MURC_HELPJ STANDARD: PRT: 449 AA.
AC 09712;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE acetylmuramoyl-L-alanine synthetase).
GN MURC OR JHP0567.
OS Helicobacter pylori 399 (Campylobacter pylori 399).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
ON NCBI_Taxid=85963;
RX SEQUENCE FROM N.A.
RX MEDLINE=9120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummimo P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the murCDF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001489; AAP06138.1; -
DR PTR: B71917; B71917; -
DR HAMAP: MF_00046; -; 1.
DR InterPro: IPR000713; Mur_Ligase.
DR InterPro: IPR004101; Mur_Ligase_C.
DR InterPro: IPR005758; MurC.
DR Pfam: PF01225; Mur_Ligase; 1.
DR Pfam: PF02875; Mur_Ligase_C; 1.
DR TIGRfams: TIGR01082; murC; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KM ATP-binding; Complete proteome.
FT NP BIND 121 127 ATP (POTENTIAL).
SQ SEQUENCE 449 AA; 50906 MW; 4EDD7CCF8A5D74F CRC64;
Query Match 76.7%; Score 33; DB 1; Length 449;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAYVPN 6
DB 180 CAIYVN 185

```

```

DE (Hepatopoeitin A).
GN HGF OR HPTA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RX SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
RA Seki T., Hagiya M., Shimomishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene.";
RL Gene 102:213-219(1991).
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RC MEDLINE=89392017; PubMed=2528952;
RA Miyazawa K., Tsudouchi H., Naka D., Takahashi K., Okigaki M.,
RA Araiaki N., Nakayama H., Hiroo S., Sakiyama O., Takahashi K.,
RA Gohda E., Daikuhara Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
RT growth factor.";
RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN [3]
RP SEQUENCE FROM N.A.
RP TISSUE=Leukocyte;
RC MEDLINE=91025062; PubMed=2145836;
RA Seki T., Ihara I., Sugimura A., Shimomishi M., Nishizawa T.,
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT "Isolation and expression of cDNA for different forms of hepatocyte
RT growth factor from human leukocyte.";
RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RC TISSUE=Liver;
RX MEDLINE=90066676; PubMed=2531289;
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimomishi M.,
RA Sugimura A., Tashiro K., Shimizu S.;
RT "Molecular cloning and expression of human hepatocyte growth factor.";
RL Nature 342:440-443(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic fibroblast;
RX MEDLINE=91334393; PubMed=1831266;
RA Weidner K.M., Araiaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA Rieder H., Fonatsch C., Tsudouchi H., Hishida T., Daikuhara Y.,
RA Birchenmeier W.;
RT "Evidence for the identity of human scatter factor and human
RT hepatocyte growth factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
RN [6]
RP SEQUENCE FROM N.A.
RA Courtney L., Elliot G., Angel S.;
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 249-695 FROM N.A.
RX MEDLINE=91369928; PubMed=1832556;
RA Miyazawa K., Kitamura A., Kitamura N.;
RT "Structural organization and the transcription initiation site of the
RT human hepatocyte growth factor gene.";
RL Biochemistry 30:9170-9176(1991).
RN [8]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=91207365; PubMed=1826837;
RA Yoshiyama Y., Araiaki N., Naka D., Takahashi K., Hiroo S., Kondo J.,
RA Nakayama H., Gohda E., Kitamura N., Tsudouchi H., Ishii T.,
RA Hishida T., Daikuhara Y.;
RT "Identification of the N-terminal residue of the heavy chain of both
RT native and recombinant human hepatocyte growth factor.";
RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-476.
RX MEDLINE=93129192; PubMed=1482348;
RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,

```

RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [10]
 RP MOTAGENESIS
 RX MEDLINE-92331602; PubMed-1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding.";
 RL EMO J. 11:2503-2510(1992).
 RN [11]
 RP STRUCTURE BY NMR OF 31-127.
 RX MEDLINE-98154323; PubMed-9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site.";
 RL Structure 6:109-116(1998).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RX MEDLINE-99036858; PubMed-9817840;
 RA Ullrich M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 RT factor at 2.0-A resolution.";
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 kringle domains.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D90334; BAA1348.1; JOINED.
 DR EMBL; D90318; BAA1348.1; JOINED.
 DR EMBL; D90319; BAA1348.1; JOINED.
 DR EMBL; D90320; BAA1348.1; JOINED.
 DR EMBL; D90322; BAA1348.1; JOINED.
 DR EMBL; D90323; BAA1348.1; JOINED.
 DR EMBL; D90324; BAA1348.1; JOINED.
 DR EMBL; D90325; BAA1348.1; JOINED.
 DR EMBL; D90326; BAA1348.1; JOINED.
 DR EMBL; D90327; BAA1348.1; JOINED.
 DR EMBL; D90328; BAA1348.1; JOINED.
 DR EMBL; D90329; BAA1348.1; JOINED.
 DR EMBL; D90330; BAA1348.1; JOINED.
 DR EMBL; D90331; BAA1348.1; JOINED.
 DR EMBL; D90332; BAA1348.1; JOINED.
 DR EMBL; D90333; BAA1348.1; JOINED.
 DR EMBL; M29145; AAS2850.1; JOINED.
 DR EMBL; M60718; AAS2648.1; JOINED.
 DR EMBL; X16323; CAA3487.1; JOINED.
 DR EMBL; M73239; AAA64297.1; JOINED.
 DR EMBL; M73240; AAA64297.1; JOINED.
 DR EMBL; AC004960; AAC71655.1; JOINED.
 DR EMBL; M75983; AAG53460.1; JOINED.
 DR EMBL; M75972; AAG53460.1; JOINED.
 DR EMBL; M75973; AAG53460.1; JOINED.
 DR EMBL; M75974; AAG53460.1; JOINED.
 DR EMBL; M75975; AAG53460.1; JOINED.
 DR EMBL; M75976; AAG53460.1; JOINED.

DR EMBL; M75977; AAG53460.1; JOINED.
 DR EMBL; M75978; AAG53460.1; JOINED.
 DR EMBL; M75979; AAG53460.1; JOINED.
 DR EMBL; M75980; AAG53460.1; JOINED.
 DR EMBL; M75981; AAG53460.1; JOINED.
 DR EMBL; M75982; AAG53460.1; JOINED.
 DR PIR; JH0579; JH0579
 DR PDB; 2HGF; 24-JUN-98.
 DR PDB; 1BHT; 18-NOV-98.
 DR PDB; 1NKT; 29-DEC-98.
 DR MEROPS; S01.976; -.
 DR GLYCOSULEDB; P14210; -.
 DR Genew; HGNC:4893; HGF.
 DR MIM; 142409; -.
 DR GO; GO:0008083; F: growth factor activity; NMS.
 DR GO; GO:0007067; P: mitosis; NMS.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRY_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR Growth factor: Kringle; Glycoprotein: Serine protease homolog; Repeat;
 KW Signal; 3D-structure: Polymorphism: Pyrolydine carboxylic acid.
 FT SIGNAL 1 31
 FT CHAIN 32 494
 FT MOD_RES 32 32
 FT DOMAIN 128 206
 FT DOMAIN 211 288
 FT DOMAIN 305 383
 FT DOMAIN 391 469
 FT DOMAIN 495 728
 FT DISULFID 70 96
 FT DISULFID 74 84
 FT DISULFID 128 206
 Query Match 76.7%; Score 33; DB 1; Length 728;
 Best Local Similarity 57.1%; Pred. No. 36;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CAVVPC 7
 Db 377 CSQIPNC 383
 ID NUGM_HYLLA STANDARD; PRT; 174 AA.
 AC Q95710;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
 GN NMD6 OR NMD6 OR NADH6
 OS Hylobates lar (Common gibbon).
 OG Mitocondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=9580;
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-Ester;
RA Arnason U., Gullberg A., Xu X.;
RT "A complete mitochondrial DNA molecule of the white-handed gibbon,
RT Hylobates lar, and comparison among individual mitochondrial genes of
RT all hominoid genera.";
RL Hereditas 124:185-189(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X89256; CAA67639.1; -
DR PIR: T11844; T11844.
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18556 MW; A956501979D20F03 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 174;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
Db 40 CAVILNC 46

RESULT 7
NU6M_PONPA STANDARD; PRT; 174 AA.
AC P92700;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
GN MTN06 OR ND6 OR NADH6.
OS Pongo pygmaeus abelii (Sumatran orangutan).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.
OX NCB1_TaxID-9601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Isolate YN93-312;
RX MEDLINE-97032590; PubMed-8875856;
RA Xu X., Arnason U.;
RT "The mitochondrial DNA molecule of Sumatran orangutan and a molecular
RT proposal for two (Bornean and Sumatran) species of orangutan.";
RL J. Mol. Evol. 43:431-437(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87707; CAA66294.1; -
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18562 MW; A3BE85628CB18FA8 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 174;
Best Local Similarity 71.4%; Pred. No. 15;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
Db 40 CAVILNC 46

RESULT 8
Y208_MYCPN STANDARD; PRT; 196 AA.
AC P75486;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG208 homolog (H10_orf196).
GN MPN291 OR MP544.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCB1_TaxID-2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-8948633;
RA Himmelfeich R., Halbert H., Plagens H., Pirkel E., Li B.-C.,
RT Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000053; AAB96192.1; -
DR PIR: S73870; S73870.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 196 AA; 22464 MW; E8E54512BC381BF CRC64;

Query Match 74.4%; Score 32; DB 1; Length 196;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVPNC 7
Db 90 CTVPNC 96

RESULT 9
HMB1_SOYBN STANDARD; PRT; 379 AA.
AC P46608;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein SBH1.
GN H1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCB1_TaxID-3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Somatic embryo;
RX MEDLINE-94169300; PubMed-7907232;
RA Ma H., McMullen M.D., Finner J.J.;
RT "Identification of a homeobox-containing gene with enhanced
RT expression during soybean (Glycine max L.) somatic embryo
RT development.";

PLant Mol. Biol. 24:465-473(1994).
 CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR INVOLVED IN EARLY
 CC EMBRYONIC DEVELOPMENT. PROBABLY BINDS TO THE DNA SEQUENCE 5'-TGAC-
 CC 3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN EMBRYONIC TISSUES. WEAKLY
 CC DETECTED IN STEMS AND HYPOCOTYL.
 CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN THE EMBRYO PROLIFERATION
 CC STAGE, INCREASES DURING EARLY SOMATIC EMBRYO DEVELOPMENT AND
 CC DECREASES THEREAFTER.
 CC -1- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L13663; AAA20882.1;
 CC PIR: S42543; S42543.
 CC HSP: P41778; IDU6.
 CC TRANSFAC: T04052;
 CC InterPro: IPR005339; ELK.
 CC InterPro: IPR001356; Homeobox.
 CC InterPro: IPR005540; KNOX1.
 CC InterPro: IPR005541; KNOX2.
 CC Pfam: PF03789; ELK; 1.
 CC Pfam: PF00046; homeobox; 1.
 CC Pfam: PF03790; KNOX1; 1.
 CC Pfam: PF03791; KNOX2; 1.
 CC ProDom: PD000010; Homeobox; 1.
 CC SMART: SM00389; HOX; 1.
 CC PROSITE: PS00027; HOMEBOX_1; 1.
 CC PROSITE: PS50071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Nuclear protein; Transcription regulation;
 CC Activator.
 CC FT DOMAIN 102 107 POLY-HIS.
 CC FT DOMAIN 109 113 POLY-ASN.
 CC FT DOMAIN 115 122 POLY-SER.
 CC FT DOMAIN 244 248 ASP/GLU-RICH (ACIDIC).
 CC FT DOMAIN 259 282 ELK DOMAIN.
 CC FT DNA_BIND 283 345 HOMEBOX (TALE-TYPE).
 CC SQ SEQUENCE 379 AA; 42374 MW; 886B020029ED5E4 CRC64;
 CC -----
 CC Query Match 74.4%; Score 32; DB 1; Length 379;
 CC Best Local Similarity 57.1%; Pred. No. 31;
 CC Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC QY 1 CAVPNC 7
 CC Db 56 CLEPNC 62
 CC -----
 CC RESULT 10
 CC PTA_HELPJ STANDARD: PRT; 519 AA.
 CC AC 092X04;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
 CC GN PTA OR JH00841.
 CC OS Helicobacter pylori J99 (Campylobacter pylori J99).
 CC OC Bacteriia; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 CC OC Helicobacteriaceae; Helicobacter.
 CC OX NCBI_TaxID=85963;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=99120557; PubMed=9923682;
 CC RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doly P.C.,
 CC Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,

Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voris G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori";
 CC Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate -> CoA + acetyl
 CC phosphate.
 CC -1- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
 CC ACETYL-COA.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE PHOSPHATE
 CC ACETYLTRANSFERASE AND BUTYRYLTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001513; AAP06419.1;
 CC PIR: D71881; D71881.
 CC InterPro: IPR004614; Pta.
 CC InterPro: IPR002505; PTA_PTB.
 CC Pfam: PF01515; PTA_PTB; 1.
 CC TRIGRAMS: TRGR00651; pta; 1.
 CC Transferase; Acyltransferase; Complete proteome.
 CC FT DOMAIN 196 519 PHOSPHATE ACETYLTRANSFERASE.
 CC SQ SEQUENCE 519 AA; 56737 MW; 3E5A71ECBBBD37 CRC64;
 CC -----
 CC Query Match 74.4%; Score 32; DB 1; Length 519;
 CC Best Local Similarity 66.7%; Pred. No. 42;
 CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 CAVPNC 6
 CC Db 361 CALPNC 366
 CC -----
 CC RESULT 11
 CC PTA_MYCTU STANDARD: PRT; 690 AA.
 CC AC P96254;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
 CC GN PTA OR RV0408 OR MT0421 OR MTGY22610.04.
 CC OS Mycobacterium tuberculosis.
 CC OC Bacteriia; Actinobacteriia; Actinomycetales;
 CC OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
 CC OX NCBI_TaxID=1773;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=H37Rv;
 CC RX MEDLINE=96295987; PubMed=9634230;
 CC RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekla E.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence";
 CC Nature 393:537-544(1998).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=CDC 1551 / Oshkosh;
 CC RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate -> CoA + acetyl
 CC phosphate.
 CC
 CC -1- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
 CC ACETYL-COA.
 CC
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, CONTAINS A DOMAIN WITH SOME
 CC SIMILARITIES WITH COH/BIOD.
 CC
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE PHOSPHATE
 CC ACETYLTRANSFERASE AND BUTYRYLTRANSFERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z84724; CAB06578.1; -
 CC EMBL: AE006946; AAK44645.1; -
 CC PIR: F70628; F70628.
 CC TIGR: MT0421; -
 CC
 CC Tuberculist: RY040414; -
 CC InterPro: IPR004614; Pta.
 CC InterPro: IPR002505; Pta_PTB.
 CC Pfam: PF01515; Pta_PTB; 1.
 CC DR Pfam: TIGR00651; Pta.1.
 CC TIGR: TIGR00651; Pta.1.
 CC KJ Transferase: Acyltransferase; Complete proteome.
 CC FT DOMAIN 365 690 PHOSPHATE ACETYLTRANSFERASE.
 CC SQ SEQUENCE 690 AA; 72948 MW; C01C412AF2810CCE CRC64;
 CC
 CC Query Match 74.4%; Score 32; DB 1; Length 690;
 CC Best Local Similarity 66.7%; Pred. No. 55;
 CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 CAVPRN 6
 CC II::II
 CC Db 530 CATHNC 535
 CC
 CC RESULT 12
 CC TR2M_AGRV4 STANDARD; PRT; 755 AA.
 CC AC P04029;
 CC DT 23-OCT-1986 (Rel. 02, Created)
 CC DT 23-OCT-1986 (Rel. 02, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
 CC GN TMS1.
 CC OS Agrobacterium tumefaciens (strain Ach5), and
 CC OS Agrobacterium tumefaciens.
 CC OG Plasmid pTiAch5, and Plasmid pTiA6NC.
 CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 CC NCBI_TaxID:176298; 358;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-Ach5; PLASMID-pTiAch5;
 CC RX MEDLINE=84207942; PubMed=6327292;
 CC RA Giesen M., de Beuckeleer M., Seurinck J., Deboeck F., de Greve H.,
 CC Lemmers M., van Montagu M., Schell J.;
 CC "The complete nucleotide sequence of the TL-DNA of the Agrobacterium
 CC tumefaciens plasmid pTiAch5.";
 CC RT EMBO J. 3:835-846(1984).
 CC RL [2]
 CC RP SEQUENCE FROM N.A.

CC PLASMID-pTiA6NC;
 CC RX MEDLINE=84170374; PubMed=6584906;
 CC RA Klee H., Montoya A., Horodyski F., Lichtenstein C., Garfinkel D.,
 CC RA Fuller S., Flores C., Peschon J., Nester E., Gordon M.;
 CC RT "Nucleotide sequence of the tms genes of the pTiA6NC octopine T1
 CC plasmid: two gene products involved in plant tumorigenesis.";
 CC RL Proc. Natl. Acad. Sci. U.S.A. 81:1728-1732(1984).
 CC
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) -> indole-3-acetamide +
 CC CO(2) + H(2)O.
 CC
 CC -1- PATHWAY: Biosynthesis of auxins from tryptophan; first step.
 CC
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF ACH5.
 CC
 CC -1- SIMILARITY: SIGNIFICANT HOMOLOG TO THE ADENINE BINDING REGION OF
 CC P-HYDROXYBENZONATE HYDROXYLASE FROM P. FLORESCENS. IT SEEMS THAT
 CC THIS PROTEIN BINDS ADENINE EITHER AS SUBSTRATE OR COFACTOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: K02554; AAA92550.1; -
 CC PIR: A04497; QOAG47.
 CC DR InterPro: IPR000759; Adnrx_reductase.
 CC DR InterPro: IPR002937; Amino_oxidase.
 CC DR InterPro: IPR006064; Glycosidase.
 CC DR InterPro: IPR000205; NAD_binding.
 CC Pfam: PF01593; Amino_oxidase; 1.
 CC DR Pfam: PF02027; Ro1b_Ro1C; 1.
 CC DR PRINTS: PR00419; ADXRPTASE.
 CC KJ T-DNA: Plasmid.
 CC FT VARIANT 718 719 NR -> IQ (IN PTA6NC).
 CC FT VARIANT 721 721 P -> A (IN PTA6NC).
 CC SQ SEQUENCE 755 AA; 83947 MW; 9FD2B83FEA0D14D CRC64;
 CC
 CC Query Match 74.4%; Score 32; DB 1; Length 755;
 CC Best Local Similarity 57.1%; Pred. No. 59;
 CC Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 CAVPNC 7
 CC II::II
 CC Db 726 CATHNC 732
 CC
 CC RESULT 13
 CC TR2M_AGRV1 STANDARD; PRT; 755 AA.
 CC AC P25017;
 CC DT 01-MAR-1992 (Rel. 21, Created)
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
 CC GN IAAM.
 CC OS Agrobacterium vitis (Rhizobium vitis).
 CC OG Plasmid pTiTM4.
 CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 CC NCBI_TaxID:373;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-TM4;
 CC RX MEDLINE=91329707; PubMed=1868204;
 CC RA Bonnard G., Vincent F., Otten L.;
 CC "Sequence of Agrobacterium tumefaciens biotype III auxin genes.";
 CC RT Plant Mol. Biol. 16:733-738(1991).
 CC RL [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-GA474;
 CC RL Otten L., de Ruifray P.;
 CC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

	CC	-1 SUBUNIT: DIMER (PROBABLE).
	CC	-1 SUBCELLULAR LOCATION: Type I membrane protein.
	CC	-1 TISSUE SPECIFICITY: FOUND IN BONE MARROW.
	CC	-1 SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.
	CC	-1 SIMILARITY: Contains 5 fibronectin type III domains.
	CC	-1 SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
	CC	
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
	CC	at the European Bioinformatics Institute. There are no restrictions on its
	CC	use by non-profit institutions as long as its content is in no way
	CC	modified and this statement is not removed. Usage by and for commercial
	CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
	CC	or send an email to license@isb-sib.ch).
	CC	-----
	DR	EMBL; M58288; AAA37673.1; -
	DR	PIR; A34898; A34898.
	DR	PDB; 1GCF; 22-OCT-97.
	DR	PDB; 1CFO; 22-OCT-97.
	DR	PDB; 1CD9; 08-MAR-00.
	DR	PDB; 1PGR; 08-MAR-00.
	DR	MGI; M61133975; Csf3r.
	DR	GO; GO:0030593; P:neutrophil chemotaxis; IMP.
	DR	InterPro; IPR002996; CR1A.
	DR	InterPro; IPR003961; FN_III.
	DR	InterPro; IPR003529; Hemtopoptn_L_F2.
	DR	Pfam; PF00041; fn3; 3.
	DR	SMART; SMO0060; FN3; 2.
	DR	PROSITE; PS01353; HEMATOPO_REC_L_F2; 1.
	KW	Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
	KW	Immunoglobulin domain; Glycoprotein; 3D-structure.
	KW	SIGNAL
	FT	CHAIN 1 26 837
	FT	DOMAIN 26 626
	FT	TRANSMEM 627 650
	FT	DOMAIN 651 837
	FT	DOMAIN 26 118
	FT	DOMAIN 122 228
	FT	DOMAIN 229 333
	FT	DOMAIN 334 431
	FT	DOMAIN 432 528
	FT	DOMAIN 529 624
	FT	DISULFID 132 143
	FT	DISULFID 249 296
	FT	DISULFID 267 310
	FT	CARBOHYD 51 51
	FT	CARBOHYD 94 94
	FT	CARBOHYD 129 129
	FT	CARBOHYD 186 186
	FT	CARBOHYD 279 279
	FT	CARBOHYD 392 392
	FT	CARBOHYD 408 408
	FT	CARBOHYD 474 474
	FT	CARBOHYD 487 487
	FT	CARBOHYD 582 582
	FT	CARBOHYD 613 613
	FT	STRAND 232 235
	FT	STRAND 249 255
	FT	STRAND 266 274
	FT	STRAND 281 286
	FT	STRAND 290 296
	FT	STRAND 303 311
	FT	STRAND 323 328
	SO	SEQUENCE 837 AA; 93406 MW; 42295E989E2C8531 CRC64;
	QY	Query Match
		Best Local Similarity 74.4%; Score 32; DB 1; Length 837;
		Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
	OY	1 CAVEPNC 7
		I :
		46 CTISPNC 52


```

RESULT 15
ID VU5_HSV7J STANDARD; PRT; 865 AA.
AC P52522;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE U5 protein.
GN U5.
OS Human herpesvirus (type 7 / strain JI) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=52278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-6 AND
CC HSV-7 U4, HSV-6 AND HSV-7 U5 AND HCMV UL27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U043400; AAC54669.1; -
CC PIR; T41909; T41909.
DR
SQ SEQUENCE 865 AA; 100187 MW; 3D3F4770B0FFAC77 CRC64;
Query Match 74.4%; Score 32; DB 1; Length 865;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CAVVPC 7
Db 199 COVERPC 205

```

Search completed: September 4, 2003, 21:01:46
 Job time : 7.07353 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 : Search time 28.7206 Seconds
(Without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-61
Perfect score: 43
Sequence: 1 CAVVPMC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	39	90.7	1704 5 094446	094446 chironomus
2	37	86.0	100 6 095KE3	095KE3 macaca fasc
3	37	86.0	409 17 09UXJ6	09UXJ6 sulfolobus
4	35	81.4	169 10 09LGO2	09LGO2 oryza sativ
5	34	79.1	24 10 08WJW1	08WJW1 zea mays (m
6	34	79.1	126 11 08BJU8	08BJU8 mus musculu
7	34	79.1	212 4 09P096	09P096 homo sapien
8	34	79.1	312 10 09LJ09	09LJ09 oryza sativ
9	34	79.1	316 4 08NCX1	08NCX1 homo sapien
10	34	79.1	368 16 053303	053303 mycobacteri
11	34	79.1	375 16 08YBN2	08YBN2 bruceella me
12	34	79.1	375 16 08FWP5	08FWP5 rhizobium su
13	34	79.1	415 16 098AG8	098AG8 rhizobium 1
14	34	79.1	433 12 091L91	091L91 white spot
15	34	79.1	446 4 08N5V5	08N5V5 homo sapien
16	34	79.1	500 11 08BLI7	08BLI7 mus musculu

17	34	79.1	558 4 08NFF0	08NFF0 homo sapien
18	34	79.1	586 4 08NFF1	08NFF1 homo sapien
19	34	79.1	605 4 08NFF2	08NFF2 homo sapien
20	34	79.1	605 11 08CCO8	08CCO8 mus musculu
21	34	79.1	644 5 061834	061834 caenorhabdi
22	34	79.1	2592 5 0810L2	0810L2 metapenaeus
23	34	76.7	73 6 09N052	09N052 macaca fasc
24	33	76.7	223 4 08MW74	08MW74 homo sapien
25	33	76.7	227 4 066S96	066S96 homo sapien
26	33	76.7	321 4 096S27	096S27 homo sapien
27	33	76.7	400 16 053697	053697 mycobacteri
28	33	76.7	400 16 08VKR4	08VKR4 mycobacteri
29	33	76.7	553 13 08UUP5	08UUP5 xenopus lae
30	33	76.7	594 5 09U3A0	09U3A0 caenorhabdi
31	33	76.7	610 5 021604	021604 caenorhabdi
32	33	76.7	610 5 077330	077330 plasmodium
33	33	76.7	751 10 08LKR4	08LKR4 arabidopsis
34	33	76.7	752 4 013597	013597 homo sapien
35	33	76.7	893 4 08N6Z0	08N6Z0 homo sapien
36	33	76.7	909 10 065027	065027 glycine max
37	33	76.7	916 8 063067	063067 glycine max
38	33	76.7	931 5 095P09	095P09 leishmania
39	33	76.7	1209 10 09SGS6	09SGS6 arabidopsis
40	33	76.7	1550 10 09LGI1	09LGI1 arabidopsis
41	33	76.7	1698 5 094438	094438 chironomus
42	33	76.7	1891 16 08E2S3	08E2S3 leptospira
43	32	74.4	130 4 08TAX3	08TAX3 homo sapien
44	32	74.4	147 4 09P0R3	09P0R3 homo sapien
45	32	74.4	157 2 09R467	09R467 agrobacteri

ALIGNMENTS

RESULT 1
094446 PRELIMINARY: PRT: 1704 AA.
ID 094446
AC 094446:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 220 kDa silk protein.
GN SP220.
OS Chironomus thummi (midge).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prelygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;
OC Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7154;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RT silk proteins sp185 and sp220."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U54641; AAA99804.1; -
DR InterPro: IPR004153; CXCXC-repeat.
DR InterPro: IPR006209; EGF-like.
DR Pfam: PF03128; CXCXC; 69.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
SQ SEQUENCE 1704 AA; 185746 MW; 3A3F20247C8F1E28 CRC64;
Query Match 90.7%; Score 39; DB 5; Length 1704;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CAVVPMC 7
DB 1442 CAVVPMC 1448
RESULT 2

095KE3 PRELIMINARY; PRT; 100 AA.

ID 095KE3

AC 095KE3

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Hypothetical 11.5 kDa protein.

OS Macaca fascicularis (Crah eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Cerebellum.

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.;

RT "Isolation of full-length cDNA clones from macaque brain cDNA

RT libraries."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB062929; BAB60725.1; -.

KM Hypothetical protein.

SQ SEQUENCE 100 AA; 11512 MW; 1AB73A899B73BEA8 CRC64;

Query Match 86.0%; Score 37; DB 6; Length 100;

Best Local Similarity 71.4%; Pred. No. 3.3;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVPNC 7

DB 36 CIVPNC 42

RESULT 3

09UXJ6 PRELIMINARY; PRT; 409 AA.

ID 09UXJ6

AC 09UXJ6

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE UDP-glucose dehydrogenase (UDP-glucose 6-dehydrogenase) (UGD)

DE (EC 1.1.1.22)

GN UGD OR SSO0810 OR ORF-C39_020.

OS Sulfobolus solfataricus.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=2287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35092 / DSM 1617 / P2;

RA MEDLINE=21332296; PubMed=11427726;

RA She Q., Singh R.K., Contaloni F., Zivanovic Y., Allard G.,

RA Aweez M.J., Chan-Weher C.C.Y., Clausen I.G., Curtis B.A.,

RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,

RA Helkamp-de Jong J., Jeffries A.C., Kozera C.J., Medina N., Peng X.,

RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,

RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,

RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;

RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."

RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

DR EMBL; Y18930; CAB57493.1; -.

DR EMBL; AE006704; AAK41109.1; -.

DR InterPro; IPR000205; NAD-binding.

DR InterPro; IPR006109; NAD_gly3P_dom.

DR InterPro; IPR001732; UDPG_MGDP_dh.

DR Pfam; PF00984; UDPG_MGDP_dh.1.

DR Pfam; PF03720; UDPG_MGDP_dh.C.1.

DR Pfam; PF03721; UDPG_MGDP_dh.N.1.

DR ProDom; PD001278; NAD_gly3P_dom.1.

KW Oxidoreductase; Complete proteome.

SQ SEQUENCE 409 AA; 44837 MW; 22EAB1AB796809D6 CRC64;

Query Match 86.0%; Score 37; DB 17; Length 409;

Best Local Similarity 71.4%; Pred. No. 12;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVPNC 7

DB 213 CIVPNC 219

RESULT 4

09LGO2 PRELIMINARY; PRT; 169 AA.

ID 09LGO2

AC 09LGO2

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE ESTs AU071032(R10630).

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Eriocaridaceae; Oryzae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(Ga3) genomic DNA, chromosome 1, PAC

RT clone:P0510F03."

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP002486; BAB03366.1; -.

DR Gramene; 09LGO2; -.

DR InterPro; IPR003657; WRKY.

DR Pfam; PF03106; WRKY.1.

DR PROSITE; PS50811; WRKY.1.

SQ SEQUENCE 169 AA; 18275 MW; DA1B876772C0A302 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 169;

Best Local Similarity 71.4%; Pred. No. 14;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVPNC 7

DB 135 CIVPNC 141

RESULT 5

08WIM1 PRELIMINARY; PRT; 24 AA.

ID 08WIM1

AC 08WIM1

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 2.7 kDa protein (fragment).

OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACAD clade; Panicoidae; Andropogoneae; Zea.

OX NCBI_TaxID=4577;

RN [1]

RP SEQUENCE FROM N.A.

RA Ananiev E.V., Lorentzen J., Bruggemann E.;

RT "Microsatellite megatracts in the maize (Zea mays L.) genome."

RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF413055; AAL55292.1; -.

DR InterPro; IPR001594; ZnF_DHHC.

DR PROSITE; PS50216; ZF_DHHC.1.

KW Hypothetical protein.

FT NON_TER 1

FT NON_TER 24

SQ SEQUENCE 24 AA; 2742 MW; 3F491B8A538ADF33 CRC64;

Query Match 79.1%; Score 34; DB 10; Length 24;

Best Local Similarity 57.1%; Pred. No. 3.6;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
11:111
Db 10 CSICPNC 16

RESULT 6

Q8BJ08 PRELIMINARY; PRT; 126 AA.
AC Q8BJ08:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Weekly similar to sodium/potassium/calcium exchanger 3 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK079049; BAC37514.1;
SQ SEQUENCE 126 AA; 14052 MW; 5F659AD717CCCD89 CRC64;

Query Match 79.1%; Score 34; DB 11; Length 126;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVNC 7
11:111
Db 35 CVTIPNC 41

RESULT 7

Q9P096 PRELIMINARY; PRT; 212 AA.
AC Q9P096:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE HSPC285 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.,
RT "Human partial CDS from cd34+ stem cells."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF161403; AAF28963.1;
FT NON_TER 1
SQ SEQUENCE 212 AA; 22078 MW; 12E1194E7A8F75D1 CRC64;

Query Match 79.1%; Score 34; DB 4; Length 212;
Best Local Similarity 71.4%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
11:111
Db 22 CAELPNC 28

RESULT 8

ID Q9LJ09 PRELIMINARY; PRT; 312 AA.
AC Q9LJ09:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to actin-like protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(CA3) genomic DNA, chromosome 1, PAC
RT clone:p0667A10.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
CC IN ALL EUKARYOTIC CELLS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
DR EMBL: AP001073; BAA89581.1;
DR Gramene; Q9LJ09;
DR InterPro: IPR004000; Actin_like.
DR Pfam: PF00022; actin; 1.
DR SMART: SM00268; ACTIN; 1.
DR Structural Protein.
KW SEQUENCE 312 AA; 34639 MW; EB7531627A3C5A09 CRC64;

Query Match 79.1%; Score 34; DB 10; Length 312;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVDPNC 7
111111
Db 29 AVDPNC 34

RESULT 9

Q8NCX1 PRELIMINARY; PRT; 316 AA.
AC Q8NCX1:
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN DKF2P434P185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Duesterhoeft A., Lauber J., Mewes H.W., Weill B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL834225; CAD38903.1;
KW Hypothetical protein.
SQ SEQUENCE 316 AA; 35263 MW; 1FA9E29F5A6DFD35 CRC64;

Query Match 79.1%; Score 34; DB 4; Length 316;
Best Local Similarity 57.1%; Pred. No. 40;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVNC 7
11:111
Db 132 CVTIPNC 138

RESULT 10

ID O53303 PRELIMINARY; PRT; 368 AA.

AC 053303;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
 DE Putative alcohol dehydrogenase (Zinc-binding dehydrogenase).
 GN ADH OR RV3086 OR MT013.07 OR MT3171.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterinae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID=1773;
 RP SEQUENCE FROM N.A.
 RC SRRAIN-H37R;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sultson J.E., Taylor K., Whitehead S., Whittam B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ernolova M.D., Salzberg S.L.,
 Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY (BY SIMILARITY)
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 CC EMBL: AL021309; CA16144.1; -;
 DR EMBL: AE007134; AKK47507.1; -;
 DR HSSP: P40394; IAGN.
 DR TIGR: MT3171; -;
 DR Tuberculist: RV3086; -;
 DR InterPro: IPR002328; Adh_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR000515; Bp_transp.
 DR InterPro: IPR002055; NAD_binding.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 DR PROSITE: PS00402; Bp_TRANSF_INN_MEMBR; 1.
 DR Oxidoreductase; zinc; Complete proteome.
 KW SEQUENCE 368 AA; 38341 MW; A03FCB6F3C8003AF CRC64;
 SO

Query Match 79.1%; Score 34; DB 16; Length 368;
 Best Local Similarity 57.1%; Pred. No. 46;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 DB 85 CSFIPNC 91

RESULT 11
 O8YBN2 PRELIMINARY; PRT; 375 AA.
 AC O8YBN2;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Alcohol dehydrogenase (EC 1.1.1.1).
 GN BME10867.
 OS Brucella melitensis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 RX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RA MEDLINE=20020109; PubMed=11756688;
 RA Delyeche V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Teleson J.-J.,
 RA Haselkorn R., Kyrides N., Overbeek R.;
 RA "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009720; AL54109.1; -;
 DR InterPro: IPR002328; Adh_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR002055; NAD_binding.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 DR Oxidoreductase; Complete proteome.
 KW SEQUENCE 375 AA; 39567 MW; D8F4009E96CDBDEA CRC64;
 SO

Query Match 79.1%; Score 34; DB 16; Length 375;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 DB 93 CVFVNC 99

RESULT 12
 O8FWP5 PRELIMINARY; PRT; 375 AA.
 AC O8FWP5;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
 DE Alcohol dehydrogenase, zinc-containing.
 GN BRA0401.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 RX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RA MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RA "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL: AE014539; AAN33598.1; -;
 DR TIGR: BRA0401; -;
 KW Complete proteome.
 SO SEQUENCE 375 AA; 39582 MW; 779A71D041C367AA CRC64;

Query Match 79.1%; Score 34; DB 16; Length 375;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 DB 93 CVFVNC 99

```

RESULT 13
Q98AG8 PRELIMINARY; PRT; 415 AA.
AC Q98AG8:
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE 8-amino-7-oxononanoate synthase.
GN MLL6006.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003008; BAB52362.1;
DR InterPro; IPR004839; AminoTransferase/2.
DR Pfam; PF00155; aminotran_1_2; 1.
KW Complete proteome.
SQ SEQUENCE 415 AA; 44108 MW; EEE219D749AE997 CRC64;

Query Match 79.1%; Score 34; DB 16; Length 415;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
DB 339 CGVPPSC 345

RESULT 14
Q91191 PRELIMINARY; PRT; 433 AA.
AC Q91191:
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ORF161 (WSV338) (WSSV394).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342572; PubMed=11448154;
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Turchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RT "The white spot syndrome virus DNA genome sequence."
RT Virology 286:7-22(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Turchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21348311; PubMed=11689662;
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus."
RT J. Virol. 75:11811-11820(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

```

```

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-Taiwan;
RX MEDLINE=20517548; PubMed=11062040;
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase."
RT Virology 277:100-110(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN-Taiwan;
RX MEDLINE=21844071; PubMed=11853398;
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells."
RT Virology 293:44-53(2002).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN-Taiwan;
RA Lo C.-F., Kou G.-H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF369029; AAK77830.1;
DR EMBL; AF332093; AAL33340.1;
DR EMBL; AF40570; AAL89262.1;
SQ SEQUENCE 433 AA; 48217 MW; 387FD4CE411ABE15 CRC64;

Query Match 79.1%; Score 34; DB 12; Length 433;
Best Local Similarity 57.1%; Pred. No. 54;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAVVNC 7
DB 348 CSLPNC 354

RESULT 15
Q08NSV5 PRELIMINARY; PRT; 446 AA.
AC Q08NSV5:
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 1200003C23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain, and Lung;
RC Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031567; AAB31567.1;
DR InterPro; IPR000132; ASX_Hydroxyl.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
SQ SEQUENCE 446 AA; 47087 MW; F2CF5E66AC083A6F CRC64;

Query Match 79.1%; Score 34; DB 4; Length 446;
Best Local Similarity 71.4%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
DB 11:1111

```

Fri Sep 5 09:16:06 2003

us-09-580-893d-61.rpt

Page 6

Db 212 CAELPNC 218

Search completed: September 4, 2003, 21:06:45
Job time : 30.815 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 41.1765 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893d-62
Perfect score: 49
Sequence: 1 CGAVVPMC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	49	100.0	8 23	AAE16727
2	49	100.0	8 23	AAE16730
3	43	87.8	8 23	AAE16716
4	43	87.8	8 23	AAE16719
5	43	87.8	91 23	ABP34373
6	40	81.6	137 22	AAU42999
7	39	79.6	395 22	AAAM25607
8	39	79.6	520 22	AAAM23901
9	39	79.6	584 22	AAAB67465

10	39	79.6	594 22	AAAG5376	Human colon cancer
11	39	79.6	748 22	AAAB60097	Human transport pr
12	39	79.6	816 22	AAAG77820	Human ion channel
13	39	79.6	816 23	AAAG61532	Human transporter
14	37	75.5	283 22	AAAE12888	Murine chordin pro
15	37	75.5	469 23	AAAB97384	Novel human protei
16	37	75.5	497 22	AAU163161	Human novel secret
17	37	75.5	497 24	AAU55430	Human novel polype
18	37	75.5	534 22	AAU15908	Human novel secret
19	37	75.5	534 22	ABU54577	Human novel polype
20	37	75.5	707 22	AAAG37118	Human polypeptide,
21	37	75.5	707 22	AAAB93234	Human protein sequ
22	37	75.5	801 22	AAE03738	Human chordin (CH
23	37	75.5	867 20	AAV01711	A human huchordin
24	37	75.5	867 20	AAAG75647	Human Tango-66/huc
25	37	75.5	885 23	AAU81856	Human PRO243.. Hom
26	37	75.5	915 23	AAAG69661	Human secreted pro
27	37	75.5	916 22	AAE03735	Murine mature chor
28	37	75.5	948 22	AAE03736	Murine chordin (CH
29	37	75.5	948 23	AAU75168	Mouse chordin poly
30	37	75.5	954 19	AAW48978	Mature human chori
31	37	75.5	954 20	AAV17821	Human PRO243 prote
32	37	75.5	954 21	AAAB01312	Human PRO243 polyp
33	37	75.5	954 23	AAU86132	Human PRO243 polyp
34	37	75.5	954 24	ABU67122	Human PRO polypept
35	37	75.5	954 24	ABU64917	Human secreted/tri
36	37	75.5	954 24	ABU60231	Human PRO polypept
37	37	75.5	954 24	ABU58351	Novel human secret
38	37	75.5	954 24	ABU55921	Human secreted/tri
39	37	75.5	954 24	ABU56302	Human secreted/tri
40	37	75.5	954 24	ABU57237	Human PRO243 prote
41	37	75.5	954 24	ABU11303	Human chordin prot
42	37	75.5	955 22	AAAE12889	Human chordin (CH
43	37	75.5	955 23	AAAG31265	Murine chordin (CH
44	37	75.5	962 22	AAE03734	Porcine IGFBP-6 tr
45	36	73.5	29 13	AAAR26998	

ALIGNMENTS

RESULT 1	
ID	AAE16727 standard; peptide; 8 AA.
XX	AAE16727:
AC	
XX	
XX	09-APR-2002 (first entry)
DE	
XX	Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
XX	
KW	Hydrolysed elastin peptide; HEP, therapy; cosmetic appearance; cardiant;
KW	vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypotensive; chronic obstructive pulmonary disease; antidiabetic; angiotensin;
KW	restenosis post-angioplasty; antidiabetic; antidiabetic; antidiabetic;
XX	
OS	Ligamentum nuchae.
OS	Synthetic.
PN	WO200191700-A2.
XX	
PD	06-DEC-2001.
XX	
XX	30-MAY-2001; 2001WO-US17384.
XX	
XX	30-MAY-2000; 2000US-0580110.
XX	30-MAY-2000; 2000US-0580156.
XX	30-MAY-2000; 2000US-0580893.
XX	30-MAY-2000; 2000US-0584001.
XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX	

PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived peptide.
 CC
 SO Sequence 8 AA;
 OY Query Match 100.0%; Score 49; DB 23; Length 8;
 DB Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 1 CGAVVPMC 8
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CGAVVPMC 8
 DB 1 CGAVVPMC 8
 RESULT 2
 AAE16730
 ID AAE16730 standard; peptide; 8 AA.
 AC
 XX AAE16730;
 DT 09-APR-2002 (first entry)
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note- "This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"
 FT Modified-site 8
 FT /note- "This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 XX
 FT
 FT
 XX
 PN W0200191700-A2.
 PD 06-DEC-2001.
 PD
 PD 30-MAY-2001; 2001WO-US17384.
 PE
 XX 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 21; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPN derived cyclic peptide.
 CC
 SO Sequence 8 AA;
 OY Query Match 100.0%; Score 49; DB 23; Length 8;
 DB Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 1 CGAVVPMC 8
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CGAVVPMC 8
 DB 1 CGAVVPMC 8
 RESULT 3
 AAE16716
 ID AAE16716 standard; peptide; 8 AA.
 AC
 XX AAE16716;
 DT 09-APR-2002 (first entry)
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..7
 FT
 FT
 XX
 PN W0200191700-A2.
 PD 06-DEC-2001.
 PD
 PD 30-MAY-2001; 2001WO-US17384.
 PE
 XX 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI WPI; 2002-106259/14.
 DR
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 XX
 PS Claim 24; Page 20; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VWPQ derived cyclic peptide.
 XX
 SQ Sequence 8 AA:
 Query Match 87.8%; Score 43; DB 23; Length 8;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
 Db 1 CGAVVPOC 8

RESULT 4
 AAE16719
 ID AAE16719 standard; peptide; 8 AA.
 XX
 AC AAE16719;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VWPQ derived peptide #13.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiac;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;
 KW restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note="This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"
 FT Modified-site 8 /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX

PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI WPI; 2002-106259/14.
 DR
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 XX
 PS Claim 24; Page 20; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VWPQ derived cyclic peptide.
 XX
 SQ Sequence 8 AA:
 Query Match 87.8%; Score 43; DB 23; Length 8;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
 Db 1 CGAVVPOC 8

RESULT 5
 ABP34373
 ID ABP34373 standard; Protein; 91 AA.
 XX
 AC ABP34373;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Human ORF3346 protein, SEQ ID NO:6692.
 XX
 KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiac; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 OS
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX

PF 24-MAY-2001; 2001WO-US17076.
 XX
 XX 24-MAY-2000; 2000US-206690P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 PI Leach MD, Shinkens RA;
 XX
 XX WPI; 2002-106200/14.
 DR N-PSDB; ABN78399.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 XX
 XX Claim 10; Page 1920; 2508pp; English.
 XX
 XX Sequences ABP31028-ABP3561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC and antiinfective activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC storage diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC disease, and infectious diseases caused by viral, bacterial, fungal
 CC and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 XX
 SO Sequence 91 AA;
 Query Match 87.8%; Score 43; DB 23; Length 91;
 Best Local Similarity 87.5%; Pred. No. 2.9;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CGAVVPNC 8
 Db 38 CGAVSPNC 45
 RESULT 6
 AAU42999
 ID AAN42999 standard; Protein; 137 AA.
 XX
 AC AAU42999;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #3895.

XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 PN
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX
 XX 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX
 XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59519.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX Example 1; SEQ ID NO 4194; 1069pp; English.
 PS
 XX
 XX Sequences AAN39105-AAN68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 137 AA;
 Query Match 81.6%; Score 40; DB 22; Length 137;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CGAVVPNC 8
 Db 45 CGAVVGNC 52
 RESULT 7
 AAM25607
 ID AAM25607 standard; Protein; 395 AA.
 XX
 AC AAM25607;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:1122.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; virocidic;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anemia;
 KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
 KW dermatological; antiallergic; antidiabetic; antidiabetic; cyostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antinaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haemotopoietic disorder; platelet disorder; asthma;
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.

XX Homo sapiens.
 OS
 XX WO200153455-A2.
 PN
 XX 26-JUL-2001.
 PD
 XX 22-DEC-2000; 2000WO-US35017.
 PF
 XX 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-048725.
 PR 25-APR-2000; 2000US-0552317.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI: 2001-457603/49.
 DR
 DR N-PSDB; AAB9548.
 DR
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 PS Claim 20; Page 232; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM251225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virocidic; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;
 CC antidiabetic; osteopathic; dermatological; antiallergic; antidiabetic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haemotopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 395 AA;

Query Match 79.68; Score 39; DB 22; Length 395;
 Best Local Similarity 75.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
 DB 169 CGIVPNC 176

RESULT 8
 AAM23901
 ID AAM23901 standard; Protein: 520 AA.

XX AAM23901;
 AC
 XX 12-OCT-2001 (first entry)

DE Rat EST encoded protein SEQ ID NO: 1426.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.

XX Rattus norvegicus.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.

XX 25-JAN-2000; 2000US-0491404.

XX 17-JUL-2000; 2000US-0617746.

XX 03-AUG-2000; 2000US-0631451.

XX 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 DR WPI: 2001-476164/51.

DR N-PSDB; AAB98560.

PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -

XX Claim 20; Page 992-993; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.

XX Sequence 520 AA;

Query Match 79.68; Score 39; DB 22; Length 520;
 Best Local Similarity 75.0%; Pred. No. 77;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
 DB 358 CGIVPNC 365

RESULT 9

AAAB67465
 ID AAB67465 standard; Protein: 584 AA.

XX AAB67465;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a calcium channel transport polypeptide.

XX Calcium channel transport polypeptide; calcium trafficking;

KW neural disorder; HIV-induced dementia; immune system disorder;
 KW rheumatoid arthritis; muscular disorder; muscle contractile dysfunction;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; arrhythmia; renal disorder;
 KW proliferative disorder; cancer; lung carcinoma; breast cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200108635-A2.
 XX
 PD 08-FEB-2001.
 XX
 PE 27-JUL-2000; 2000WO-US20392.
 XX
 PR 28-JUL-1999; 99US-0145958.
 PR 18-AUG-1999; 99US-0149446.
 PR 14-MAR-2000; 2000US-0189064.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, NI J, Shi Y;
 XX
 DR WPI: 2001-138604/14.
 DR N-PSDB: AAF55042.
 XX
 PT New isolated nucleic acid useful for diagnosing, detecting, or treating
 PT or preventing diseases associated with anomalies in calcium trafficking
 PT across the plasma membrane -
 XX
 PS Claim 11; Page 256-257; 259pp; English.
 XX
 CC The present sequence represents a calcium channel transport polypeptide.
 CC The polynucleotides, polypeptides, and antibodies are useful for
 CC preventing, treating, or ameliorating diseases associated with anomalies
 CC in calcium trafficking across the plasma membrane. They are used to
 CC diagnose, detect and treat or prevent diseases or conditions such as
 CC neural disorders (e.g. HIV-induced dementia), immune system disorders
 CC (e.g. rheumatoid arthritis), muscular disorders (e.g. muscle contractile
 CC dysfunction), reproductive disorders, gastrointestinal disorders,
 CC pulmonary disorders, cardiovascular disorders (e.g. arrhythmias), renal
 CC disorders, proliferative disorders, and/or cancerous diseases and
 CC conditions (e.g. lung carcinoma or breast cancer).
 CC
 SQ Sequence 584 AA:
 QY
 Query Match 79.6%; Score 39; DB 22; Length 584;
 Best Local Similarity 75.0%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 1 CGAVVPMC 8
 358 CGIVPNC 365
 RESULT 10
 AAG75376
 ID AAG75376 standard; Protein; 594 AA.
 XX
 AC AAG75376;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:6140.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 PD 05-APR-2001.
 XX

PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI: 2001-235357/24.
 DR N-PSDB: AAH34781.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 7588-7590; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytosolic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC present invention.
 CC N.B. Pages 666 to 662 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 CC
 SQ Sequence 594 AA:
 QY
 Query Match 79.6%; Score 39; DB 22; Length 594;
 Best Local Similarity 75.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 1 CGAVVPMC 8
 368 CGIVPNC 375
 RESULT 11
 AAB60097
 ID AAB60097 standard; Protein; 748 AA.
 XX
 AC AAB60097;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human transport protein TPPT-17.
 XX
 KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200078953-A2.
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16668.
 XX
 PR 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.

```

PR 28-OCT-1999; 99US-0162287.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N,
PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI: 2001-041424/05.
DR N-PSDB: AAF27117.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
XX Claim 2; Page 118-120; 165pp; English.
XX
XX The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
XX
XX Sequence 748 AA;
SO
Query Match 79.6%; Score 39; DB 22; Length 748;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 522 CGIVFPMC 529

RESULT 12
AAG77820
ID AAG77820 standard; Protein: 816 AA.
AC
AC AAG77820;
XX
XX 05-DEC-2001 (first entry)
DE Human ion channel 23927 protein (IC23927).
XX
XX IC23927; human; ion channel 23927; gene therapy; screening assay;
KW predictive medicine; pharmacogenetics; pain disorder;
KW central nervous system disorder; ion transport; ion conductance;
KW membrane bound protein; membrane excitability; membrane polarisation;
KW synaptic transmission; pain signalling; cell activation; cell growth;
KW cell proliferation; cell differentiation; cell migration;
KW muscle contraction.
XX
XX Homo sapiens.
XX
XX Location/Qualifiers
FH Key 114..128
FT Domain /note="Transmembrane domain 1"
FT 146..168
FT Domain /note="Transmembrane domain 2"
FT 178..195
FT Domain /note="Transmembrane domain 3"
FT 199..210
FT Domain /note="Transmembrane domain 4"
FT 233..254
FT Domain /note="Transmembrane domain 5"
FT 269..287
FT Domain /note="Pore domain: an overall hydrophobic domain"
FT 298..320
FT Domain /note="Transmembrane domain 6"
FT 445..465
FT Domain /note="Transmembrane domain 7"
FT 482..502
FT Domain /note="Transmembrane domain 8"
FT 510..532

```

```

FT /note="Transmembrane domain 9"
FT 539..554
FT Domain /note="Transmembrane domain 10"
FT 570..594
FT Domain /note="Transmembrane domain 11"
FT 637..653
FT Domain /note="Pore domain: an overall hydrophobic domain"
FT 666..687
FT /note="Transmembrane domain 12"
XX
XX W0200164881-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06530.
XX
XX 29-FEB-2000; 2000US-0185938.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Curtis RAJ, Slios-santiago I;
PI
XX WPI: 2001-570694/64.
DR N-PSDB: AAF78802, AAF78803.
XX
XX Isolated ion channel polypeptide IC23927 useful in screening assays and
PT treatment of disorders, e.g., central nervous system disorders and pain
PT disorders -
XX
XX Claim 14; Fig 1; 142pp; English.
XX
XX The present sequence represents the human ion channel 23927 (IC23927)
CC protein, which is claimed in the invention. The ion channel family of
CC proteins is a large family of membrane bound proteins responsible for a
CC wide range of transport and signalling functions in cells. The invention
CC comprises IC23927 proteins and nucleic acids. The IC23927 nucleic acids
CC and proteins may be used in screening assays, predictive medicine (e.g.
CC diagnostic assays and pharmacogenetics) and treatment of disorders
CC characterised by insufficient production of IC23927 (e.g. central nervous
CC system disorders, pain disorders, or disorders of cellular growth,
CC differentiation or migration). The IC23927 nucleic acids and proteins are
CC also useful as targets for developing modulating agents to regulate a
CC variety of cellular processes, such as: ion transport (e.g. ion
CC conductance); membrane excitability and/or polarisation; synaptic
CC transmission (e.g. pain signalling); cell activation; cell proliferation;
CC cell growth; cell differentiation; cell migration and muscle contraction.
CC Administration of a modulator of IC23927 (especially by gene therapy) may
CC be used to treat pain or a pain disorder.
XX
XX Sequence 816 AA;
SQ
Query Match 79.6%; Score 39; DB 22; Length 816;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 590 CGIVFPMC 597

RESULT 13
ABG61532
ID ABG61532 standard; Protein: 816 AA.
AC
AC ABG61532;
XX
XX 27-AUG-2002 (first entry)
DE Human transporter and ion channel, TRICH2, Incyte ID 2907828CD1.
XX
XX Human; transporter and ion channel; TRICH2; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;

```

cell proliferative disorder; cervical cancer; breast cancer;
 neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
 myotonic dystrophy; cataplexia; endocrine disorder; diabetes;
 Grave's disease; gastrointestinal disorder; Crohn's disease;
 renal disorder; good pasture's syndrome; viral infection; cirrhosis;
 bacterial infection; fungal infection; parasitic infection;
 protozoal infection; helminthic infection; cardiovascular disorder;
 atherosclerosis; hepatic disease.

Homo sapiens.

W0200240541-A2.

23-MAR-2002.

25-OCT-2001; 2001WO-US46055.

27-OCT-2000; 2000US-243989P.
 03-NOV-2000; 2000US-245904P.
 09-NOV-2000; 2000US-247673P.
 17-NOV-2000; 2000US-249661P.
 20-NOV-2000; 2000US-252232P.
 01-DEC-2000; 2000US-250790P.

(INCY-) INCYTE GENOMICS INC.

Tang YF, Yue H, Nguyen DB, Hafalla AJA, Elliott VS, Lu Y;
 Walla NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;
 Rankumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzal Y, Khan FA;
 Thangaveilu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;
 Ison CH, Das D, Raumann BE, Policky JL, Kearney L.

WPI: 2002-463570/49.
 N-PSDB; ABR83211.

New transporters and ion channels (TRICH) polypeptides, useful for
 diagnosing, preventing, and treating disorders associated with an
 abnormal expression or activity of TRICH, e.g. immunological, muscular
 or renal disorders

Claim 1; Page 132-134; 178pp; English.

The invention relates to human transporters and ion channels (TRICH)
 polypeptides, a naturally occurring amino acid sequence 90 % identical to
 TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 a recombinant polynucleotide comprising a promoter sequence operably
 linked to the TRICH polynucleotide, a cell transformed with the
 recombinant polynucleotide, a transgenic organism comprising the
 TRICH, and screening for compounds which bind to TRICH, modulate
 TRICH, modulate TRICH expression or are ant/agonists of TRICH.
 The polypeptides are useful for diagnosing, treating, and
 preventing transport, neurological, muscle, immunological disorders
 (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 proliferative disorders such as cancers (e.g. leukaemia, cervical or
 breast cancer), neurodegenerative disorders (e.g. Parkinson's disease,
 Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 cataplexia), endocrine disorders (e.g. diabetes, Grave's disease),
 gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 (e.g. good pasture's syndrome), viral, bacterial, fungal, parasitic,
 protozoal and helminthic infections, cardiovascular disorders (e.g.
 atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 other diseases and disorders detailed in the specification. They can also
 be used in assessing the effects of exogenous compounds on the
 expression of nucleic acid and amino acid sequences of transporters and
 ion channels. TRICH or its fragments may also be used in screening for
 compounds that specifically bind to and modulate the activity of TRICH.
 The polynucleotides can be used to create knock-in humanized animals or
 transgenic animals to model human disease. The present sequence
 represents a TRICH protein.

Sequence 816 AA;

Query Match 79.6%; Score 39; DB 23; Length 816;
 Best Local Similarity 75.0%; Pred. No. 12e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVNC 8
 II | III
 Db 590 CGIVPNC 597

RESULT 14
 AAEI2888
 ID AAEI2888 standard; Protein; 283 AA.
 XX AAEI2888;
 AC
 XX
 XX
 DT 15-JAN-2002 (first entry)
 XX
 XX
 DE Murine chordin protein.
 XX
 XX Murine; chordin-like-2 protein; CHL-2; hypotensive; cardiant; cytostatic;
 XX Rheumatoid arthritis; cachexia; arrhythmia; osteoporosis; hepatitis;
 XX hypertension; myocardial infarction; stomach cancer; ulcer; gene therapy;
 XX antiinflammatory; tissue regeneration; osteoporosis; muscular dystrophy;
 XX congestive heart failure; hepatoma; angina.

OS Mus musculus.
 XX
 XX W0200164885-A1.
 XX
 XX
 PD 07-SEP-2001.
 XX
 XX
 PF 02-MAR-2001; 2001WO-US06891.
 XX
 XX
 PR 02-MAR-2000; 2000US-186462P.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Zhang K, Linh C, Nakayama N;
 XX
 XX WPI: 2001-648246/74.

Novel Chordin-like-2 polypeptides useful for diagnosing and treating
 diseases and conditions affecting bone density, tissue regeneration,
 skeletal muscle, heart, stomach and liver, and for wound healing

Example 1; Fig 2; 167pp; English.

The invention relates to chordin-like-2 (CHL-2) proteins and their
 corresponding nucleic acid molecules. The invention also provides
 pharmaceutical compositions and methods for the diagnosis, treatment,
 amelioration, and/or prevention of diseases, disorders and conditions
 associated with CHL2 proteins. Chordin-like-2 (CHL-2) proteins are useful
 for diagnosing and treating diseases and conditions affecting bone
 density (osteoporosis, osteoporosis, osteoarthritis and rheumatoid
 arthritis), skeletal muscle (e.g., cachexia and muscular dystrophy),
 heart (arrhythmia, angina, hypertension, myocardial infarction and
 congestive heart failure), stomach (stomach cancer and ulcer) and
 liver (hepatitis and hepatoma), for promoting tissue regeneration
 and wound healing, for ex vivo expansion of haematopoietic stem cells
 and gene therapy performed through such cells. CHL2 protein is also
 useful for identifying CHL2 protein receptors. A selective binding
 agent (SBA) or fragment is useful for detecting or quantitating the
 amount of CHL2 protein. A transgenic non-human mammal is useful for
 determining whether a compound inhibits CHL2 protein activity or
 CHL2 protein production. The present sequence is murine chordin protein.

Sequence 283 AA;

Query Match 75.5%; Score 37; DB 22; Length 283;
 Best Local Similarity 75.0%; Pred. No. 98;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
 ||| ||:|
 Db 232 CGAGVPHC 239

RESULT 15

ABB97384
 ID ABB97384 standard; Protein: 469 AA.

AC ABB97384;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 652.

KM Human: antihaemic; vulnerary; antiinflammatory; immunomodulator;
 KM antiinfertility; cerebroprotective; cyostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

OS Homo sapiens.

PN M0200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-292408/33.

PT An isolated polynucleotide for treating diseases associated with its

PS Claim 20; SEQ ID NO 652; 509pp; English.

CC The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat

CC stroke and cancer, to screen for drugs, to treat inflammatory conditions

CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention.

SQ Sequence 469 AA;

Query Match 75.5%; Score 37; DB 23; Length 469;
 Best Local Similarity 85.7%; Pred. NO. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GAVVPNC 8
 ||:|||||
 Db 389 GSVVPNC 395

Search completed: September 4, 2003, 21:00:35
 Job time : 42.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 12.4706 Seconds
(Without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-62

Perfect score: 49

Sequence: 1 CGAVPNC 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	38	77.6	167	1 AG0369	ferredoxin-type pr
2	37	75.5	469	2 T46929	hypothetical prote
3	37	75.5	469	2 T46930	hypothetical prote
4	36	73.5	61	2 AH2852	conserved hypothet
5	36	73.5	61	2 F97629	hypothetical prote
6	36	73.5	74	2 C23734	insulin-like growt
7	36	73.5	90	2 B86560	9 kDa-Cysteine-ric
8	36	73.5	90	2 A72064	cysteine rich oute
9	36	73.5	99	2 S47084	lipid transfer lik
10	36	73.5	105	1 FE1WT	ferredoxin [3Fe-4S
11	36	73.5	123	2 D75287	ferredoxin - Deino
12	36	73.5	240	2 A39842	insulin-like growt
13	36	73.5	837	2 F96561	unknown protein [l
14	35	71.4	109	2 F71644	ferredoxin (fdaA)
15	35	71.4	116	2 C97860	ferredoxin [import
16	35	71.4	118	1 PSN33M	phospholipase A2 (
17	35	71.4	118	1 PSN33B	phospholipase A2 (
18	35	71.4	250	2 G72495	probable polysulfi
19	35	71.4	329	2 T17033	leucine rich repea
20	35	71.4	330	2 J02262	polygalacturonase
21	35	71.4	336	2 S15997	chitinase (BC 3.2.
22	35	71.4	341	2 T05764	hypothetical prote
23	35	71.4	506	2 T35923	probable tRNA synt
24	35	71.4	532	2 T47335	hypothetical prote
25	35	71.4	572	2 T27869	sphingomyelin phos
26	35	71.4	579	2 JC7629	membrane-type friz
27	35	71.4	605	2 T15291	sphingomyelin phos
28	35	71.4	612	2 C90374	hypothetical prote
29	35	71.4	653	2 T25194	hypothetical prote

30	34	69.4	92	2 T04395	probable phospholi
31	34	69.4	129	2 B35216	FP14 protein - low
32	34	69.4	196	2 T46525	probable transcrip
33	34	69.4	368	2 H69335	iron-sulfur cluste
34	34	69.4	369	2 F69407	iron-sulfur cluste
35	34	69.4	893	2 T17276	hypothetical prote
36	34	69.4	921	2 S49965	probable membrane
37	34	69.4	1013	2 T10659	probable serine/th
38	34	69.4	1029	2 T00712	protein kinase hom
39	34	69.4	1188	2 D86236	protein F14N23.5 l
40	34	69.4	1244	2 S37034	DNA-directed DNA p
41	41	67.3	81	2 D69477	hypothetical prote
42	33	67.3	102	2 S04126	hypothetical prote
43	33	67.3	107	2 F82278	ferredoxin VC0804
44	33	67.3	226	2 JN0464	insulin-like growt
45	33	67.3	241	2 G87530	endonuclease III f

ALIGNMENTS

RESULT 1

AG0369 ferredoxin-type protein NapF [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 30-Sep-2002

C:Accession: AG0369

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrcl

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0369

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-167 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92282.1; PID:q15980993; GSPDB:GN00175

C:Genetics:

A:Gene: napF

C:Superfamily: ferredoxin protein NapF; ferredoxin [4Fe-4S] homology

Query Match 77.6%; Score 38; DB 1; Length 167;

Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8

Db 146 CGACVPGC 153

RESULT 2

T46929 hypothetical protein DKFZp434E2220.1 - human

C:Species: Homo sapiens (man)

C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: T46929

R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Well, B.; Wiemann, S.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24136

A:Accession: T46929

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-469 <AA>

A:Cross-references: EMBL:AL157433

A:Experimental source: adult testis; clone DKFZp434E2220

C:Genetics:

A>Note: DKFZp434E2220.1

Query Match 75.5%; Score 37; DB 2; Length 469;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVVPNC 8
DB 389 GSVVPNC 395

RESULT 3

T46930

hypothetical protein DKFZP43400420.1 - human

C:Species: Homo sapiens (man)

C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: T46930

R:Ottewald, B.; Obermaier, B.; Mewes, H.W.; Mell, B.; Wiemann, S.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z4136

A:Accession: T46930

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-469 <AAA>

A:Cross-references: EMBL:AL157434

A:Experimental source: adult testis; clone DKFZP43400420

C:Genetics:

A:Note: DKFZP43400420.1

Query Match

Best Local Similarity 75.5%; Score 37; DB 2; Length 469;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVVPNC 8

DB 389 GSVVPNC 395

RESULT 4

AH2852

conserved hypothetical protein Atu2249 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AH2852

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yeo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH2852

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-61 <KUR>

A:Cross-references: GB:AE008688; PIDN:AA43238.1; PID:q17740722; GSPDB:GN00186

C:Genetics:

A:Gene: Atu2249

A:Map position: circular chromosome

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 61;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8

DB 6 CGTIVPGC 13

RESULT 5

F97629

hypothetical protein AGR_C_4091 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: F97629

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouellet, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: F97629

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-61 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87991.1; PID:q15157403; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_4091

A:Map position: circular chromosome

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 61;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8

DB 6 CGTIVPGC 13

RESULT 6

C23734

Insulin-like growth factor-binding protein 6 - pig (fragments)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 11-Jan-2000

C:Accession: C23734

R:Shimasaki, S.; Gao, L.; Shimonaka, M.; Ling, N.

Mol. Endocrinol. 5, 938-948, 1991

A:Title: Isolation and molecular cloning of insulin-like growth factor-binding protein

A:Reference number: A23734; MUID:92049376; PMID:1719383

A:Accession: C23734

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-74 <SHD>

C:Superfamily: insulin-like growth factor binding protein 1; thyroglobulin type-I rep

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 74;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8

DB 30 CGVTPNC 37

RESULT 7

B86560

9 kDa-Cysteine-rich lipoprotein [imported] - Chlamydomonas reinhardtii (strain J138)

C:Species: Chlamydomonas reinhardtii

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: B86560

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydomonas reinhardtii J138.

A:Reference number: A86491; MUID:20303493; PMID:10871362

A:Accession: B86560

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <STO>

A:Cross-references: GB:BA000008; MUID:98978929; PIDN:BA98764.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: omca

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 90;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8

DB 55 CGSVTPNC 62

```

RESULT 8
A:72064
Cysteine rich outer membrane protein 3 CP0193 [imported] - Chlamydia pneumoniae (str
N:Alternate names: 9 kDa cysteine-rich lipoprotein
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: A72064; G81603
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:992060606; PMID:10192388
A:Accession: A72064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <ARN>
A:Cross-references: GB:AE001640; GB:AE00163; NID:94376845; PIDN:AA018698.1; PID:9437685
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: G81603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <REA>
A:Cross-references: GB:AE002180; GB:AE002161; NID:97189117; PIDN:AAF38066.1; PID:9718912
C:Genetics:
A:Gene: omca; CP0193

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 90;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 55 CGSVYPSC 62

RESULT 9
A:547084
Lipid transfer like protein - cowpea
C:Species: Vigna unguiculata (cowpea)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Sep-1997
C:Accession: S47084
R:Krause, A.; Sigrist, C.J.A.; Dehning, I.; Sommer, H.; Broughton, W.G.
submitted to the EMBL Data Library, June 1994
A:Description: Accumulation of transcript encoding a lipid transfer-like protein during
A:Reference number: S47084
A:Accession: S47084
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-99 <KRA>
A:Cross-references: EMBL:X79604; NID:9499033; PID:9499034

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 99;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 92 CGVYPNC 99

RESULT 10
A:FETW
Ferredoxin [3Fe-4S][4Fe-4S] - Thermus aquaticus (tentative sequence)
C:Species: Thermus aquaticus
C:Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 03-Nov-2000
C:Accession: A00216
R:Sato, S.; Nakazawa, K.; Hon-Nami, K.; Oshima, T.

```

```

Biochim. Biophys. Acta 668, 277-289, 1981
A:Title: Purification, some properties and amino acid sequence of Thermus thermophilu
A:Reference number: A00636; MUID:81184605; PMID:7225412
A:Accession: A00216
A:Molecule type: Protein
A:Residues: 1-69;97-105 <SAT>
A:Experimental source: strain HB8; ATCC 27634
R:Hille, R.; Yoshida, T.; Tarr, G.E.; Williams Jr., C.H.; Ludwig, M.I.; Fee, J.A.; Ke
J. Biol. Chem. 258, 13008-13013, 1983
A:Title: Studies of the ferredoxin from Thermus thermophilus.
A:Reference number: A92402; MUID:84032522; PMID:6313685
A:Contents: annotation; composition
A:Note: we have positioned residues 70-96 by homology with other Azotobacter-type fer
C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
C:Keywords: 3Fe-4S; 4Fe-4S; duplication; electron transfer; iron-sulfur protein; meta
F.16,49/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted
F.20,39,42,45/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match
Best Local Similarity 73.5%; Score 36; DB 1; Length 105;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 42 CGACVPAC 49

RESULT 11
A:D75287
Ferredoxin - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: D75287
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <WHT>
A:Cross-references: GB:AE002064; GB:AE000513; NID:96660134; PIDN:AAF11876.1; PID:9646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2330
A:Map position: 1
C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 123;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 88 CGACVPAC 95

RESULT 12
A:A39842
Insulin-like growth factor-binding protein 6 precursor - human
N:Alternate names: IGFBP-6; insulin-like growth factor II-binding protein
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_revision 31-Dec-1995 #text_change 03-Dec-1999
C:Accession: A39842; B23734; PH0144; A35470; C35803
R:Kiefer, M.C.; Maslarczyk, F.R.; Bauer, D.M.; Zapf, J.
J. Biol. Chem. 266, 9043-9049, 1991
A:Title: Identification and molecular cloning of two new 30-kDa insulin-like growth f
A:Reference number: A39842; MUID:91225006; PMID:1709161
A:Accession: A39842
A:Molecule type: mRNA

```

A:Residues: 1-240 <KTE>
 A:Cross-references: GB:M62402; NID:g184813; PIDN:AAB06187.1; PID:g184814
 R:Shimasaki, S.; Gao, L.; Shimonaka, M.; Ling, N.
 Mol. Endocrinol. 5, 938-948, 1991
 A:Title: Isolation and molecular cloning of insulin-like growth factor-binding protein-6
 A:Reference number: A23734; MUID:92049376; PMID:1719383
 A:Accession: B23734
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 'C', 3-240 <SHI>
 A:Cross-references: GB:M69054; NID:g183893; PIDN:AA88070.1; PID:g183894
 R:Adress, D.L.; Birnbaum, R.S.
 Biochem. Biophys. Res. Commun. 176, 213-218, 1991
 A:Title: A novel human insulin-like growth factor binding protein secreted by osteoblast
 A:Reference number: PH0143; MUID:91207395; PMID:1850257
 A:Accession: PH0144
 A:Molecule type: protein
 A:Residues: 'XX', 30-31, 'P', 33-39, 'X', 41-42 <AND>
 A:Experimental source: cell line V-2
 R:Martin, J.L.; Willetts, K.E.; Baxter, R.C.
 J. Biol. Chem. 265, 4124-4130, 1990
 A:Title: Purification and properties of a novel insulin-like growth factor-II binding pr
 A:Reference number: A35470; MUID:90154107; PMID:2154495
 A:Accession: A35470
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 28, 'A', 30-39 <MAR>
 R:Zapf, J.; Klefer, M.; Merryweather, J.; Maslarz, F.; Bauer, D.; Born, W.; Fischer, J.A
 J. Biol. Chem. 265, 14892-14898, 1990
 A:Title: Isolation from adult human serum of four insulin-like growth factor (IGF) bindi
 ic tumor hypoglycemia.
 A:Reference number: A35803; MUID:90368661; PMID:1697583
 A:Accession: C35803
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 'AA', 30-31, 'H', 33-54, 'QXG' <ZAP>
 C:Genetics:
 A:Gene: GDB:IGFBP6
 A:Cross-references: GDB:127456; OMIM:146735
 A:Map position: 12pter-12qter
 C:Superfamily: insulin-like growth factor binding protein 1; thyroglobulin type I repeat
 C:Keywords: glycoprotein
 F:1-27/Domain: signal sequence #status predicted <SIC>
 F:28-240/Product: insulin-like growth factor-binding protein 6 #status predicted <MAT>
 F:163-234/Domain: thyroglobulin type I repeat homology <THY1>
 F:229/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 240;
 Best Local Similarity 62.5%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 II III
 Db 71 CGVYTPNC 78

RESULT 13
 F96561
 unknown protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F96561
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Matzli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F96561
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-837 <STO>
 A:Cross-references: GB:AE005173; NID:g6850339; PIDN:AAF29402.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F915.1
 A:Map position: 1

Query Match 73.5%; Score 36; DB 2; Length 837;
 Best Local Similarity 62.5%; Pred. No. 96;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 I IIII
 Db 469 CMTVPNC 476

RESULT 14
 F71644
 ferredoxin (fdx) RP829 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: F71644
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slichteritz-Ponten, T.; Alsmark
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: F71644
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-109 <AND>
 A:Cross-references: GB:A735273; GB:A735269; NID:g3861237; PIDN:CA15254.1; PID:g386
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: fdx, RP829
 A:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
 F:2-58/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 71.4%; Score 35; DB 2; Length 109;
 Best Local Similarity 62.5%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 II III
 Db 43 CGVVPNC 50

RESULT 15
 C97860
 ferredoxin [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: C97860
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: C97860
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-116 <KUD>
 A:Cross-references: GB:AE006914; PIDN:AAU03821.1; PID:g15620421; GSPDB:GN00173
 A:Genetics:
 A:Gene: fdx
 C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

Query Match 71.4%; Score 35; DB 2; Length 116;
 Best Local Similarity 62.5%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 II III

Db 47 CGVCVPDC 54

Search completed: September 4, 2003, 21:10:52
Job time : 13.4706 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 : Search time 6.94118 Seconds

(Without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893d-62

Perfect score: 49

Sequence: 1 CGAVVPNC 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	75.5	948	1	CHRD_MOUSE
2	37	75.5	955	1	CHRD_HUMAN
3	36	73.5	78	1	FER_THER
4	36	73.5	90	1	OM3_CHLPN
5	36	73.5	99	1	NLTP_VIGUN
6	35	71.4	240	1	IBP6_HUMAN
7	35	71.4	108	1	FER_RICPR
8	35	71.4	118	1	PA23_NAJMO
9	35	71.4	118	1	PA23_NAJNG
10	35	71.4	330	1	PG12_ARATH
11	35	71.4	330	1	PG1P_PIRCO
12	35	71.4	336	1	CH12_ORYSA
13	35	71.4	485	1	MKR4_HUMAN
14	34	69.4	129	1	V061_FOPPV
15	34	69.4	715	1	SYM_METWA
16	34	69.4	921	1	YIB7_YEAST
17	34	69.4	1207	1	DPOL_ASEB7
18	34	69.4	1244	1	DPOL_ASEB6
19	33	67.3	68	1	NLTP2_PRIUR
20	33	67.3	102	1	NLTP2_HORVU
21	33	67.3	152	1	PA23_ORPHA
22	33	67.3	152	1	PA2H_LARSE
23	33	67.3	226	1	IBP6_RAT
24	33	67.3	287	1	YIIV_ECOLI
25	33	67.3	304	1	IBP2_RAT
26	33	67.3	305	1	IBP2_MOUSE
27	33	67.3	486	1	EHSL_SCHPO
28	33	67.3	712	1	SYM_METAC
29	33	67.3	894	1	CYAS_HUMAN
30	33	65.3	54	1	FER_PEPAS
31	32	65.3	55	1	FER_CLOAC
32	32	65.3	55	1	FER_CLOSP
33	32	65.3	55	1	FER_CLOSP

34	32	65.3	55	1	FER_CLOST	PB0168 clostridium
35	32	65.3	60	1	FER1_CHILI	P00204 chlorobium
36	32	65.3	61	1	FER1_CHLIE	ORKE66 chlorobium
37	32	65.3	61	1	FER1_CHLIT	P00205 chlorobium
38	32	65.3	207	1	CLDB_HUMAN	O75308 homo sapien
39	32	65.3	238	1	DEOD_ECOLI	P09743 escherichia
40	32	65.3	249	1	TSG_DROME	P54356 drosophila
41	32	65.3	260	1	YAMP_RHOCA	P14172 rhodobacter
42	32	65.3	277	1	MCR4_ECOLI	P24200 escherichia
43	32	65.3	318	1	CH11_ORYSA	P24626 oryza sativ
44	32	65.3	342	1	PG11_PHAVU	P35334 phaseolus v
45	32	65.3	353	1	PSBA_MEDSA	P04998 medicago sa

ALIGNMENTS

RESULT 1	CHRD_MOUSE	STANDARD	PRT	948 AA.
AC	0920E2:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Chordin precursor.			
GN	CHRD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-BESUL/EL:			
RA	Lu B., Bachiller D., Agius E., Piccolo S., De Robertis E.M.;			
RT	"Bmp-binding domains in the chordin secreted protein.";			
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-99000848: PubMed-9782094;			
RA	Pappano W.N., Scott I.C., Clark T.G., Eddy R.L., Shows T.B.;			
RA	Greenspan D.S.;			
RT	"Coding sequence and expression patterns of mouse chordin and mapping of the cognate mouse chrd and human CHRD genes.";			
RL	Genomics 52:236-239(1998).			
CC	-1- FUNCTION: Dorsalizing factor. Key developmental protein that dorsalizes early vertebrate embryonic tissues by binding to ventralizing TGF-beta family bone morphogenetic proteins (Bmps) and sequestering them in latent complexes.			
CC	-1- SUBCELLULAR LOCATION: Secreted (By similarity).			
CC	-1- DEVELOPMENTAL STAGE: Detected at high levels of a in 7 dpc mouse embryos; its level decrease at later developmental stages and in adult tissues.			
CC	-1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.			
CC	-1- SIMILARITY: Contains 4 WFC domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: AF069276: A019895.1; -			
DR	EMBL: AF069501: AAC68867.1; -			
DR	MGD: MGI:131326; Chrd.			
DR	GO: GO:0007420; P:brain development; IMP.			
DR	GO: GO:0007389; P:pattern specification; IMP.			
DR	GO: GO:0001501; P:skeletal development; IMP.			
DR	InterPro: IPR006559; SOG.			
DR	InterPro: IPR001007; WVF_C.			
DR	Pfam: PF00093; WVF_4.			
DR	SMART: SM00566; SOG; 3.			

Query Match	Best Local Similarity	Score 37;	DB 1;	Length 948;
Matches 6;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1 CGAATPNC 8			
DB	897 CGAGVPHC 904			
<p>RESULT 2</p> <p>CHRD_HUMAN STANDARD; PRT; 955 AA.</p> <p>Q9H2X0; Q95254; Q9H2D3; Q9H2W6; Q9P0Z2; Q9P0Z3; Q9P0Z4; Q9P0Z5;</p> <p>16-OCT-2001 (Rel. 40, Created)</p> <p>16-OCT-2001 (Rel. 40, Last sequence update)</p> <p>28-FEB-2003 (Rel. 41, Last annotation update)</p> <p>Chordin precursor.</p> <p>GN CHRD.</p> <p>OS Homo sapiens (Human).</p> <p>CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.</p> <p>NCBI_Taxid=9606;</p> <p>11</p> <p>SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).</p> <p>MEDLINE=21366001; PubMed=11472837;</p> <p>Miller C., Lemaire P., Orselli B., Guglielmi P., Francois V.;</p> <p>"The human chordin gene encodes several differentially expressed spliced variants with distinct BMP opposing activities.";</p> <p>Mech. Dev. 106:85-96(2001).</p> <p>12</p> <p>SEQUENCE OF 115-955 FROM N.A. (ISOFORM 5).</p> <p>MEDLINE=99000848; PubMed=9782094;</p> <p>Pappano W.N., Scott I.C., Clark T.G., Eddy R.L., Shows T.B., Greenspan D.S.;</p> <p>"Coding sequence and expression patterns of mouse chordin and mapping of the cognate mouse chrd and human CHRD genes.";</p> <p>Genomics 52:236-239(1998).</p> <p>13</p> <p>SEQUENCE OF 51-125; 705-762; 784-850 AND 872-932 FROM N.A. Lu B., Bachiller D., Agius E., Larrain J., Piccolo S., Nieters A., De Robertis E.M.;</p> <p>"Bmp binding modules in the chordin patterning protein.";</p> <p>Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.</p> <p>1-1 FUNCTION: Dorsalizing factor. Key developmental protein that dorsalizes early vertebrate embryonic tissues by binding to ventralizing TGF-beta family bone morphogenetic proteins (BMPs) and sequestering them in latent complexes (By similarity).</p> <p>1-1 SUBCELLULAR LOCATION: Secreted (By similarity).</p> <p>1-1 ALTERNATIVE PRODUCTS:</p> <p>Event=Alternative splicing; Named isoforms=5;</p> <p>Comment=Experimental confirmation may be lacking for some isoforms;</p> <p>Name=1;</p> <p>Isoid=Q9H2X0-1; Sequence=Displayed;</p> <p>Name=2;</p> <p>Isoid=Q9H2X0-2; Sequence=VSP_001069, VSP_001070;</p> <p>Name=3;</p>				

```

CC CC      IsoId-Q9H2X0-3; Sequence=VSP_0010171, VSP_0010172;
CC CC      Name=4;
CC CC      IsoId-Q9H2X0-4; Sequence=VSP_0010173, VSP_0010174;
CC CC      Name=5;
CC CC      IsoId-Q9H2X0-5; Sequence=VSP_0010175;
CC CC      -1- TISSUE SPECIFICITY: Expressed at the highest level in liver.
CC CC      -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
CC CC      -1- SIMILARITY: Contains 4 WMC domains.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; WMC_1; 2.
CC CC      DR PROSITE; PS01084; WMC_2; 2.
CC CC      KW developmental protein; Repeat; glycoprotein; signal;
CC CC      Alternative splicing.
CC CC      EMBL, AF2099928; AAC35767.1; -
CC CC      DR EMBL, AF2099929; AAC35768.1; -
CC CC      DR EMBL, AF2099930; AAC35769.1; -
CC CC      DR EMBL, AF076612; AAC69835.1; -
CC CC      DR EMBL, AF283325; AAC35784.1; -
CC CC      DR EMBL, AF136632; AAF70236.1; -
CC CC      DR EMBL, AF136633; AAF70237.1; -
CC CC      DR EMBL, AF136634; AAF70238.1; -
CC CC      DR EMBL, AF136635; AAF70239.1; -
CC CC      DR Genew; HGNC:1949; CHRD..
CC CC      DR MIM; 603475; -
CC CC      DR GO: GO:0001501; P:skeletal development; TAS.
CC CC      DR InterPro: IPR001007; WMC_C.
CC CC      DR Pfam; PF00093; WMC; 4.
CC CC      DR PROSITE; PS01208; W
```

```

RESULT 3
FER_THERM
ID FER_THERM STANDARD: PRT: 78 AA.
AC P03942.
DT 23-OCT-1986 (Rel. 02, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ferredoxin.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP STRAIN-HB8 / ATCC 27634;
RC MEDLINE=81184605; PubMed=7225412;
RX Sato S., Nakazawa K., Hon-Nami K., Oshima T.;
RA "Purification, some properties and amino acid sequence of Thermus
thermophilus HB8 ferredoxin.";
RL Biochim. Biophys. Acta 668:277-289(1981).
[2]
RN RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.64 ANGSTROMS).
RC STRAIN-HB8 / ATCC 27634;
RX MEDLINE=21537789; PubMed=11681700;
RA Macedo-Ribeiro S., Martins B.M., Pereira P.J., Buse G., Huber R.,
RA Soullame T.;
RT "New insights into the thermostability of bacterial ferredoxins:
RT high-resolution crystal structure of the seven-iron ferredoxin from
RT Thermus thermophilus.";
RL J. Biol. Inorg. Chem. 6:663-674(2001).
[3]
RN RP COMPOSITION.
RC STRAIN-ATCC 696;
RX MEDLINE=84032522; PubMed=6313685;
RA Hille R., Yoshida T., Tarr G.E., Williams C.H. Jr., Ludwig M.I.,
RA Fee J.A., Kent T.A., Hyngh B.H., Munc E.;
RT "Studies of the ferredoxin from Thermus thermophilus.";
RL J. Biol. Chem. 258:13008-13013(1983).
CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER AND A 3FE-4S CLUSTER.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR InterPro: IPR000813; 7Fe_ferredoxin.
DR Pfam: PF00037; fer4.1.
DR PRINTS: PR00354; 7FE8SRDOXIN.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
KW Electron transport; Iron-sulfur; Repeat; 4Fe-4S; 3Fe-4S; 3D-structure.
FT METAL 8 8 IRON-SULFUR 1 (3FE-4S).
FT METAL 16 16 IRON-SULFUR 2 (4FE-4S).
FT METAL 20 20 IRON-SULFUR 2 (4FE-4S).
FT METAL 39 39 IRON-SULFUR 2 (4FE-4S).
FT METAL 42 42 IRON-SULFUR 2 (4FE-4S).
FT METAL 45 45 IRON-SULFUR 2 (4FE-4S).
FT METAL 49 49 IRON-SULFUR 1 (3FE-4S).
FT CONFLICT 6 6 E->O (IN REF.1).
SQ SEQUENCE 78 AA; 8687 MW; 12F54B3069BC4FC0 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 78;
Best Local Similarity 75.0%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

AC Q92725; Q9J015;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cysteine-rich outer membrane protein 3 precursor (9 kDa-cysteine-rich
DE lipoprotein) (9KD-CRP).
DE OMCA OR CPN0558 OR CP0193.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gail S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Moyn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN-J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto K., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: THIS OUTER MEMBRANE PROTEIN IS ASSOCIATED WITH
CC DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES
CC (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER
CC ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor.
CC -1- DEVELOPMENTAL STAGE: EXTRACELLULAR ELEMENTARY BODY STAGE
CC SPECIFIC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE001640; AAD18698.1; -
DR EMBL: AE002180; AAF38066.1; -
DR EMBL: AP002547; BAA8764.1; -
DR PIR: A72064; A72064.
DR PIR: B86560; B86560.
DR TIGR: CP0193; -
DR InterPro: IPR003517; Chlam_OMP3.
DR Pfam: PF03503; Chlam_OMP3; 1.
DR PRINTS: PR01335; CHLAMIDIOM3.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 90 CYSTEINE-RICH OUTER MEMBRANE PROTEIN 3.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 90 AA; 9410 MW; 5B0372524B86260B CRC64;

Query Match 73.5%; Score 36; DB 1; Length 90;

```


RL FEBS Lett. 255:253-258(1989).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE-98241382; PubMed-9572875;
 RA Neumann G.M., Marino J.A., Bach L.A.;
 RT "Identification of O-glycosylation sites and partial characterization
 of carbohydrate structure and disulfide linkages of human
 insulin-like growth factor binding protein 6.";
 RL Biochemistry 37:6572-6585(1998).
 RN [8]
 RP DISULFIDE BONDS.
 RX MEDLINE-99262603; PubMed-10329650;
 RA Neumann G.M., Bach L.A.;
 RT "The N-terminal disulfide linkages of human insulin-like growth
 factor-binding protein-6 (hIGFBP-6) and hIGFBP-1 are different as
 determined by mass spectrometry.";
 RL J. Biol. Chem. 274:14587-14594(1999).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PM: O-LINKED GLYCANS CONSIST OF HEX (PROBABLY GAL), HEXNAC
 (PROBABLY GALNAC) AND SIALIC ACID RESIDUES. MAJOR GLYCOPHORS
 CONSIST OF 8-16 MONOSACCHARIDES (BY SIMILARITY TO IGFBP-6
 EXPRESSED RECOMBINANTLY IN CHO CELLS).
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 thyroglobulin type-I domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M62402; AAB06187.1; -
 DR EMBL; AJ006952; CA07346.1; -
 DR EMBL; BC003507; AAH03507.1; -
 DR EMBL; BC005007; AAH05007.1; -
 DR EMBL; BC010162; AAH10162.1; -
 DR EMBL; BC011708; AAH11708.1; -
 DR EMBL; M69054; AAH8070.1; -
 DR PIR; A39842; A39842.
 DR HSSP; P24593; 1BOE.
 DR Genew; HGNC:5475; IGFBP6.
 DR MIM; 146735; -
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR00716; Thyroglobulin_1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00086; thyroglobulin_1; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00211; TY; 1.
 DR PROSITE; PS00222; IGF_BINDING; FALSE_NEG.
 DR PROSITE; PS00484; THYROGLOBULIN_1; 1.
 KW Growth factor binding: Signal: Glycoprotein: Polymorphism.
 FT SIGNAL 1 24
 FT CHAIN 25 240
 FT DOMAIN 186 234
 FT DISULFID 29 32
 FT DISULFID 40 44
 FT DISULFID 57 63
 FT DISULFID 71 84
 FT DISULFID 78 104
 FT DISULFID 163 190
 FT DISULFID 201 212
 FT DISULFID 214 234
 FT CARBOHYD 126 126
 FT CARBOHYD 144 144
 FT CARBOHYD 145 145
 O-LINKED (HEXNAC. . .) (BY SIMILARITY).
 O-LINKED (HEXNAC. . .) (BY SIMILARITY).
 O-LINKED (HEXNAC. . .) (BY SIMILARITY).

FT CARBOHYD 146 146 O-LINKED (HEXNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 152 152 O-LINKED (HEXNAC. . .) (BY SIMILARITY).
 FT VARIANT 236 236 T->P (IN dbSNP:1053134).
 FT CONFLICT 2 2 T->C (IN REF. 4).
 FT SEQUENCE 240 AA; 25322 MW; 285308231C025009 CRC64;
 SO
 Query Match 73.5%; Score 36; DB 1; Length 240;
 Best Local Similarity 62.5%; Pred. No. 8.9;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CGAVVPC 8
 DB 71 CGVTPNC 78
 RESULT 7
 FER_RICPR STANDARD; PRT; 108 AA.
 ID FER_RICPR
 AC O92C08;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ferredoxin.
 GN FDXA OR RP829.
 OS Rickettsia prowazekii.
 OC Bacteria: Proteobacteria: Alphaproteobacteria: Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=782;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE-99039499; PubMed-9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sierheltz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria.";
 RL Nature 396:133-140(1998).
 CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
 ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS (BY
 SIMILARITY).
 CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER AND A 3FE-4S CLUSTER (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AJ235273; CA15254.1; -
 DR PIR; F71644; F71644.
 DR HSSP; P00214; 1FRH.
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR InterPro; IPR00813; 7Fe_ferredoxin.
 DR Pfam; PF00037; Fer4_2.
 DR PRINTS; PR00354; 7FE8SFEROXIN.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
 KW Electron transport; Iron-sulfur; Repeat; 4Fe-4S; 3Fe-4S;
 FT SIGNAL 1 1
 FT CHAIN 2 2
 FT INIT_MET 0 0
 FT METAL 8 8
 FT METAL 16 16 BY SIMILARITY.
 FT METAL 20 20 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).
 FT METAL 39 39 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 42 42 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 45 45 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 49 49 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).
 FT METAL 49 49 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).
 SO SEQUENCE 108 AA; 12434 MW; 412EB4465A23D59 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 108;
 Best Local Similarity 62.5%; Pred. No. 6.9;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
 DB 42 CGVCPDC 49

RESULT 8

PA23_NAJMO STANDARD; PRT; 118 AA.

AC P00604;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phospholipase A2 isozyme CM-III (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase).
 OS Naja mossambica (Mozambique cobra).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidea; Elapidae; Elapinae; Naja.
 OC NCBI_TaxID=8644;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=77222126; PubMed=880314;
 RA Joubert F.J.;
 RT "Naja mossambica mossambica venom. Purification, some properties and the amino acid sequences of three phospholipases A (CM-I, CM-II and CM-III)."
 RT Biochim. Biophys. Acta 493:216-227(1977).
 CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY. GROUP I SUBFAMILY.
 CC PIR; A00744; PSNJ3M.
 DR HSSP; P00598; IPOA.
 DR InterPro: IPR001211; PhospholipaseA2.
 DR Pfam: PF00068; Phoslip; 1
 DR PRINTS; PR00389; PPHLIPASE2.
 DR ProDom; PD000303; PhospholipaseA2; 1.
 DR SMART; SM00085; PA2c; 1.
 DR PROSITE; PS00119; PA2_ASP; 1.
 DR PROSITE; PS00118; PA2_HIS; 1.
 KW Hydrolyase; Lipid degradation; Calcium; Multigene family.
 FT ACT_SITE 47 47
 FT ACT_SITE 92 92
 FT DISULFID 11 70
 FT DISULFID 26 117
 FT DISULFID 28 44
 FT DISULFID 43 98
 FT DISULFID 50 91
 FT DISULFID 59 84
 FT DISULFID 77 89
 FT METAL 27 27
 FT METAL 29 29
 FT METAL 31 31
 FT METAL 48 48
 FT SEQUENCE 118 AA; 13302 MW; 0CF2A6A2AAC11BC2 CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 118;
 Best Local Similarity 75.0%; Pred. No. 7.4;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
 DB 84 CGAVPNC 91

RESULT 9

PA23_NAJNG STANDARD; PRT; 118 AA.

AC P00605;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phospholipase A2 isozyme III (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase).
 OS Naja nigricollis (Black-necked spitting cobra).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidea; Elapidae; Elapinae; Naja.
 OC NCBI_TaxID=8654;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RA Baker D.;
 RT "Structure and function of snake venom toxins."
 RL (in) Walter R., Meleishofer J. (eds.);
 RL Peptides: chemistry, structure and function, pp.17-30, Ann Arbor Science, Ann Arbor (1975).
 CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: LD(50) is 0.63 mg/Kg by intravenous injection.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY. GROUP I SUBFAMILY.
 CC HSSP; P00598; IPOA.
 DR InterPro: IPR001211; PhospholipaseA2.
 DR Pfam; PF00068; Phoslip; 1.
 DR PRINTS; PR00389; PPHLIPASE2.
 DR ProDom; PD000303; PhospholipaseA2; 1.
 DR SMART; SM00085; PA2c; 1.
 DR PROSITE; PS00119; PA2_ASP; 1.
 DR PROSITE; PS00118; PA2_HIS; 1.
 KW Hydrolyase; Lipid degradation; Calcium.
 FT ACT_SITE 47 47
 FT ACT_SITE 92 92
 FT DISULFID 11 70
 FT DISULFID 26 117
 FT DISULFID 28 44
 FT DISULFID 43 98
 FT DISULFID 50 91
 FT DISULFID 59 84
 FT DISULFID 77 89
 FT METAL 27 27
 FT METAL 29 29
 FT METAL 31 31
 FT METAL 48 48
 FT SEQUENCE 118 AA; 13268 MW; F6F2ACA8A4C11BCF CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 118;
 Best Local Similarity 75.0%; Pred. No. 7.4;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 10

PC12_ARATH STANDARD: PRT: 330 AA.
 ID PG12_ARATH
 AC Q9M5J8; Q9FL58;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Polygalacturonase inhibitor 2 precursor (Polygalacturonase-inhibiting protein) (PGIP-2).
 DE PGIP2 OR AT5G06870 OR MOJ9.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneo T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miya]ima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence features of the regions of 1,381,565 bp covered by twenty one physically assigned P1 and TAC clones";
 RL DNA Res. 5:131-145(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RATLs) sequenced by the SSP consortium (Sally/Stanford/PGEC).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 5-330 FROM N.A.
 RA Park B., Jin Y., Nam S., Kim H.;
 RT "Arabidopsis thaliana polygalacturonase inhibiting protein 2 (PGIP2) gene";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an important factor for plant resistance to phytopathogenic fungi.
 CC -1- SUBCELLULAR LOCATION: Cell-wall associated.
 CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITTING PROTEIN FAMILY.
 CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AB010697; BAB11145.1; -
 DR EMBL: AY035121; AAK59626.1; -
 DR EMBL: AF229250; AAF69828.1; -
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_plant.
 DR Pfam: PF00560; LRR: 3.
 DR PRINTS: PR00019; LEURICHRPT.
 DR Signal: Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.
 KM SIGNAL
 FT CHAIN 1 21 POTENTIAL.
 FT CHAIN 2 22 330 POLYGALACTURONASE INHIBITOR 2.
 FT REPEAT 118 141 LRR 1.
 FT REPEAT 142 166 LRR 2.
 FT REPEAT 167 192 LRR 3.
 FT REPEAT 194 215 LRR 4.
 FT REPEAT 261 285 LRR 5.
 FT REPEAT 287 308 LRR 6.
 FT DISULFID 25 55 BY SIMILARITY.
 FT DISULFID 56 63 BY SIMILARITY.
 FT DISULFID 298 320 BY SIMILARITY.
 FT DISULFID 322 329 BY SIMILARITY.
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 330 AA: 37067 MW; 9911DF015DDC280 CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 330;
 Best local similarity 62.5%; Pred. No. 18;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAVVPMC 8
 DB 322 CGAPLSPC 329
 RESULT 11
 PCIP_PYRCO STANDARD: PRT: 330 AA.
 ID PGIP_PYRCO
 AC Q05091;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Polygalacturonase inhibitor precursor (Polygalacturonase-inhibiting protein).
 DE PGIP.
 GN Pyrus communis (Pear).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids I; Rosales; Rosaceae; Maloideae; Pyrus.
 OX NCBI_TaxID=23211;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bartlett;
 RX MEDLINE=94151422; PubMed=8108494;
 RA Stoltz H.U., Powell A.L., Damon S.E., Greve L.C., Bennett A.B.,
 RA Labavitch J.M.;
 RT "Molecular characterization of a polygalacturonase inhibitor from Pyrus communis L. cv Bartlett.";
 RL Plant Physiol. 102:133-138(1993).
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=cv. Bartlett; TISSUE=Fruit;
 RA Abu-Goukh A.A., Greve L.C., Labavitch J.M.;
 RT "Purification and partial characterization of 'Bartlett' pear polygalacturonase inhibitors";
 RL Physiol. Mol. Plant Pathol. 23:111-122(1983).
 CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an important factor for plant resistance to phytopathogenic fungi.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITTING PROTEIN FAMILY.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L09264; AAA33865.1; -
 DR PIR: JQ2262; JQ2262.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_plant.
 DR Pfam: PF00560; LRR: 3.
 DR PRINTS: PR00019; LEURICHRPT.
 DR Signal: Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.
 KM SIGNAL
 FT CHAIN 1 24 POTENTIAL.
 FT CHAIN 2 25 330 POLYGALACTURONASE INHIBITOR.
 FT REPEAT 118 141 LRR 1.
 FT REPEAT 142 166 LRR 2.
 FT REPEAT 167 192 LRR 3.

```

FT REPEAT 194 215 LRR 4.
FT REPEAT 217 237 LRR 5.
FT REPEAT 261 285 LRR 6.
FT REPEAT 287 309 LRR 7.
FT DISULFID 27 57 BY SIMILARITY.
FT DISULFID 58 65 BY SIMILARITY.
FT DISULFID 298 320 BY SIMILARITY.
FT DISULFID 322 329 BY SIMILARITY.
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 330 AA; 36504 MW; 36BA2BCF420393E3 CRC64;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 330;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 322 CGAPLPS 329

RESULT 12
CH12_ORYSA STANDARD; PRT; 336 AA.
ID CH12_ORYSA
AC P25765;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Basic endochitinase 2 precursor (EC 3.2.1.14).
GN RCH10.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatloideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91238706; PubMed=2034221;
RA Zhu Q., Lamb C.J.;
RT "Isolation and characterization of a rice gene encoding a basic
RT chitinase."
RL Mol. Gen. Genet. 226:289-296(1991).
CC -1- FUNCTION: THIS PROTEIN FUNCTIONS AS A DEFENSE AGAINST CHITIN
CC CONTAINING FUNGAL PATHOGENS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS 1A WHICH IS SIMILAR TO
CC CLASS 1B IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL
CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL
CC HYDROLASES).
CC PIR: S15997; S15997.
DR HSP: P23951; 2BA.
DR Gramene: P23765;
DR InterPro: IPR001002; Chitin_binding_1.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; Chitin_bind_1; 1.
DR Pfam: PF00182; Glyco_hydro_19; 1.
DR PRINTS: PR00451; CHITINBINDNG.
DR ProDom: PD000609; Chitin_binding_1; 1.
DR ProDom: PD354900; Glyco_hydro_19; 1.
DR SMART: SM00270; ChitBD1; 1.
DR PROSITE: PS000026; CHITIN_BINDING; 1.
DR PROSITE: PS00773; CHITINASE_19_1; 1.
DR PROSITE: PS00774; CHITINASE_19_2; 1.
DR PROSITE: PS00773; CHITINASE_19_2; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Multigene family.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 336 BASIC ENDOCHITINASE 2.

```

```

FT DOMAIN 22 64 CHITIN-BINDING (BY SIMILARITY).
FT DISULFID 24 39 BY SIMILARITY.
FT DISULFID 33 45 BY SIMILARITY.
FT DISULFID 38 52 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
FT DISULFID 99 161 BY SIMILARITY.
FT DISULFID 99 161 BY SIMILARITY.
FT DISULFID 174 182 BY SIMILARITY.
FT DISULFID 281 313 BY SIMILARITY.
SQ SEQUENCE 336 AA; 35565 MW; 7F4EC126265DEA84 CRC64;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 336;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAVPNC 8
Db 30 GAVPNC 36

RESULT 13
MKR4_HUMAN STANDARD; PRT; 485 AA.
ID MKR4_HUMAN
AC 013434;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Makorin 4 (zinc finger protein 127-Xp) (ZNF127-Xp).
GN MKR4 OR ZNF127L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hendrich B.D., Longstreet M., Gustashaw K., Nicholls R.D.;
RT "An X-linked homologue of the autosomal imprinted gene ZNF127
RT escapes X chromosome inactivation."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBA databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 4 C3H1-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U41315; AAA99070.1; -.
CC Genew: HGNC:7115; MKR4.
CC InterPro: IPR000571; Znf_CCCH.
CC InterPro: IPR001841; Znf_fing.
CC Pfam: PF00097; Zf-C3HC4; 1.
CC Pfam: PF00642; Zf-CCCH; 4.
CC SMART: SM00184; RING; 1.
CC SMART: SM00356; Znf_C3H1; 4.
CC PROSITE: PS00518; ZF_RING_1; 1.
CC PROSITE: PS00519; ZF_RING_2; 1.
KW Zinc-finger; Repeat.
FT ZN_FING 96 114 C3H1-TYPE 1.
FT ZN_FING 125 143 C3H1-TYPE 2.
FT ZN_FING 249 267 C3H1-TYPE 3.
FT DOMAIN 271 298 MAKORIN-TYPE CYS-HIS.
FT ZN_FING 316 370 RING-TYPE.
FT ZN_FING 405 425 C3H1-TYPE 4.
SQ SEQUENCE 485 AA; 52909 MW; AE28B962344CFE CRC64;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 485;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```



```

OY 1 CGAVVPMC 8
DB 214 CGRTVPSC 221

RESULT 14
V061_F0MPV STANDARD; PRT: 129 AA.
ID V061_F0MPV STANDARD; PRT: 129 AA.
AC P21972;
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative CC-type chemokine FPV061.
GN FPV061 OR FP14.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Altopoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP-1;
RX MEDLINE=90324937; PubMed=2165135.
RA Tartaglia J., Winslow J., Goebel S., Johnson G.P., Taylor J.,
RA Paolletti E.;
RT "Nucleotide sequence analysis of a 10.5 kbp HindIII fragment of
RT fowlpox virus: relatedness to the central portion of the vaccinia
RT virus HindIII D region.";
RL J. Gen. Virol. 71:1517-1524(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: LOW SIMILARITY TO THE INTERCRINE BETA FAMILY (SMALL
CC CTOKINE C-C) (CHEMOKINE CC).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: X17202; CAA35065.1;
CC DR EMBL: AF198100; AAF44405.1;
CC DR PIR: B35216; B35216.
CC KW Hypothetical protein; Cytokine.
CC SQ SEQUENCE 129 AA; 14110 MW; EDC25E79875E40DB CRC64;

Query Match 69.4%; Score 34; DB 1; Length 129;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 GAVVPMC 8
DB 87 GAIFPMC 93

RESULT 15
SYM_METMA STANDARD; PRT: 715 AA.
ID SYM_METMA STANDARD; PRT: 715 AA.
AC O8BY14;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (Meters).
GN METG OR MM0867.
OS Methanosarcina mazel (Methanosarcina frisia).
OS Archaea; Euryarchaeota; Euryarchaeota orders Incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.

```

```

OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierer A., Baumer S., Jacobl C.,
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA Bhattacharya A., Lyklis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazel: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -1- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA(fMet) aminoacylation (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC Metg subfamily 1.
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: AE013312; AAM30563.1; ALT_INIT.
CC DR HAMAP: MF_00098; fused; 1.
CC DR InterPro: IPR004495; MetG_Cterm.
CC DR InterPro: IPR001412; tRNA-synt_1.
CC DR InterPro: IPR002304; tRNA-synt_met.
CC DR InterPro: IPR002547; tRNA_bind.
CC DR Pfam: PF01588; tRNA_bind; 1.
CC DR PIRSF: PIRSF001528; MetGrs_dimerising; 1.
CC DR PRINTS: PRO1041; TRNASTNTMET.
CC DR TIGRFAMs: TIGR00398; metG; 1.
CC DR TIGRFAMs: TIGR00399; metG_C_term; 1.
CC DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
CC DR PROSITE: PS50886; TRBD; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
CC FT SITE 20 30
CC FT SITE 334 338
CC FT DOMAIN 613 715
CC FT METAL 151 151 TRNA-BINDING.
CC FT METAL 154 154 ZINC (BY SIMILARITY).
CC FT METAL 163 163 ZINC (BY SIMILARITY).
CC FT METAL 163 163 ZINC (BY SIMILARITY).
CC FT METAL 167 167 ZINC (BY SIMILARITY).
CC FT BINDING 337 337 ATP (BY SIMILARITY).
CC SQ SEQUENCE 715 AA; 79964 MW; 926A7E57855DA9CD CRC64;

Query Match 69.4%; Score 34; DB 1; Length 715;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPMC 8
DB 473 CGOVVYMC 480

Search completed: September 4, 2003, 21:01:47
Job time : 7.94118 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 32.8235 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-62

Perfect score: 49

Sequence: 1 CGAVVPC 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp-mhc:*
8: sp-organella:*
9: sp-phage:*
10: sp_plant:*
11: sp-rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	87.8	532	17	Q8TNS9
2	39	79.6	362	16	Q9RD10
3	39	79.6	857	4	Q9ULQ1
4	38	77.6	89	10	Q41378
5	38	77.6	167	16	Q8ZCP1
6	38	77.6	304	3	Q8XOP1
7	37	75.5	61	16	Q92N61
8	37	75.5	83	16	Q98HT2
9	37	75.5	413	4	Q8N2W7
10	37	75.5	469	4	Q9NS53
11	37	75.5	469	4	Q9NS52
12	37	75.5	707	4	Q8N2K5
13	37	75.5	707	4	Q8N2K5
14	36	73.5	29	6	Q96SK7
15	36	73.5	61	16	Q9TRY1
16	36	73.5	90	2	Q46183

17	36	73.5	111	4	Q9H2B5
18	36	73.5	123	16	Q9RR29
19	36	73.5	390	17	Q8TWP9
20	36	73.5	697	5	Q9GZ35
21	36	73.5	837	10	Q9M819
22	36	73.5	872	10	Q8H140
23	35	71.4	890	10	Q8VYS2
24	35	71.4	67	10	P82900
25	35	71.4	67	10	P82901
26	35	71.4	95	2	Q9ZAZ1
27	35	71.4	96	10	Q9FER9
28	35	71.4	100	10	Q8H8Y6
29	35	71.4	100	10	Q8H8Y4
30	35	71.4	101	10	Q8H4W8
31	35	71.4	116	16	Q92G41
32	35	71.4	134	5	Q9U7C7
33	35	71.4	176	12	Q9DGM8
34	35	71.4	232	16	Q932N8
35	35	71.4	250	17	Q9Y8M7
36	35	71.4	253	10	Q9LUX1
37	35	71.4	327	10	Q38695
38	35	71.4	329	10	Q38738
39	35	71.4	330	10	P93270
40	35	71.4	330	10	Q9LKC8
41	35	71.4	330	10	Q22522
42	35	71.4	341	10	Q81843
43	35	71.4	506	16	Q9ZBH7
44	35	71.4	532	10	Q9M190
45	35	71.4	564	5	Q10916

ALIGNMENTS

RESULT 1
Q8TNS9 PRELIMINARY; PRT: 532 AA.
AC Q8TNS9: 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cellulosomal protein.
OS MA2203.
OS Methanosarcina acetivorans.
OC Archaea: Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endlitz M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smitrov S., Alnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL, AE010906; AA005599.1; -;
KW Complete proteome.
SQ SEQUENCE 532 AA: 60558 MW: 3A12ADFAD67B43 CRC64;
Query Match 87.8%; Score 43; DB 17; Length 532;
Best Local Similarity 87.5%; Pred. No. 3.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 172 CGAVPNC 179

RESULT 2

Q9RD10 PRELIMINARY; PRT; 362 AA.
AC Q9RD10; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
GN Hypothetical protein SC00901.
OS SC00901 OR SC01.34C.
OC Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteriales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145.
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S., Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL39107; CAB62692.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 362 AA; 38401 MW; 9119D6F0C3F34DF CRC64;

Query Match 79.6%; Score 39; DB 16; Length 362;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 43 CGAVPNC 50

RESULT 3
Q9UL01 PRELIMINARY; PRT; 857 AA.

AC Q9UL01; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
GN KIAA1169 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirose M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis from size-fractionated cDNA libraries from human brain."
RL DNA Biol. 6:329-336(1999).
DR EMBL; AB032995; BAA86483.2;
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001220; Lectin_legB.

DR InterPro; IPR005820; M+channel_nlg.
DR Pfam; PF00520; Ion_trans_2.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
KW Ionic Channel; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 857 AA; 98609 MW; 439B63104579C992 CRC64;

Query Match 79.6%; Score 39; DB 4; Length 857;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 631 CGAVPNC 638

RESULT 4

Q41378 PRELIMINARY; PRT; 89 AA.
AC Q41378; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
GN Lipid transfer protein (Fragment).
OS Senecio odoratus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteroideae; Asteraceae; Asteroidae;
OX NCBI_TaxID=34237;
RN [1]
RP SEQUENCE FROM N.A.
RA Pyee J., Kolatukudy P.E.;
RT "Epidermis-specific transcripts including one that encodes a new class of lipid transfer proteins in Kleinia odora.";
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; L53791; AAA3933.1;
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 89 AA; 9218 MW; FC20EB89543516C6 CRC64;

Query Match 77.6%; Score 38; DB 10; Length 89;
Best Local Similarity 62.5%; Pred. No. 5.9;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 82 CGAVPNC 89

RESULT 5

Q8ZCF1 PRELIMINARY; PRT; 167 AA.
AC Q8ZCF1; 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
GN Ferredoxin-type protein NapF.
OS NAPF OR YPO3040 OR Y1442.
OC Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360.
RA Parkhill J., Wren B.W., Thomson N.R., Tlball R.W., Holden M.T.G., Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L., Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M., Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moutle S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague".
 RL Nature 413:523-527(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-KIM5 / Biovar Mediaevalis;
 RX MEDLINE=2137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.,
 RT "Genome sequence of *Yersinia pestis* KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL: AJ414155; CAC92282.1;
 DR EMBL: AE013747; AM85014.1;
 DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
 DR InterPro: IPR004496; NapF.
 DR Pfam: PF00037; fer4; 3.
 DR TIGRPFAMs: TIGR00402; napF; 1.
 DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 3.
 KW Complete proteome.
 SQ SEQUENCE 167 AA; 18075 MW; E2B5594D08E8CE6C CRC64;

Query Match 77.6%; Score 38; DB 16; Length 167;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
 |||
 Db 146 CGACVPGC 153

RESULT 6
 O8XOF1 PRELIMINARY; PRT; 304 AA.
 AC O8XOF1;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Related to putative nucleotide binding protein (NUPB).
 GN B14A6.050.
 OS Neurospora crassa.
 OC Eukaryota: Fungi: Ascomycota: Pezizomycotina: Sordariomycetes;
 OC Sordariales: Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Alpn V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL670007; CAD21307.1;
 SQ SEQUENCE 304 AA; 32237 MW; 2108A23CEF2B0E56 CRC64;

Query Match 77.6%; Score 38; DB 3; Length 304;
 Best Local Similarity 75.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
 |||
 Db 195 CGFVCPNC 202

RESULT 7
 ID 092N61 PRELIMINARY; PRT; 61 AA.
 AC 092N61;

DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein R02361.
 GN R02361 OR SMC02699.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria: Proteobacteria: Alphaproteobacteria: Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masuy D.,
 RA Pohl T., Portetelle D., Puenhler A., Purnelle B., Ranspaeper U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591790; CAC46940.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 61 AA; 7006 MW; A1F9EF3435054001 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 61;
 Best Local Similarity 62.5%; Pred. No. 6.5;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
 |||
 Db 6 CGSLVPGC 13

RESULT 8
 Q98HT2 PRELIMINARY; PRT; 83 AA.
 ID Q98HT2;
 AC Q98HT2;
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein msr2723.
 GN MSR2723.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria: Proteobacteria: Alphaproteobacteria: Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneke T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003000; BAB49784.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 83 AA; 9748 MW; DB612DDCF7DFD51 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 83;
 Best Local Similarity 62.5%; Pred. No. 8.7;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
 |||
 Db 28 CGSLVPGC 35

RESULT 9

```

08N2W7
ID 08N2W7 PRELIMINARY; PRT; 413 AA.
AC 08N2W7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC029879; AAH29879.1; -.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00093; VWC; 3.
DR SMART: SM00214; VWC; 3.
DR PROSITE: PS01208; VWF_C; 2.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 413 AA; 43822 MW; FCFD20584BEE0E06 CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 413;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 GAVVNC 8
DB 360 GAVVNC 367

RESULT 10
O9NSS3 PRELIMINARY; PRT; 469 AA.
AC O9NSS3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Well B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ductus deferens adenocarcinoma;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL157433; CAB75657.1; -.
DR EMBL: BC017704; AAH17704.1; -.
DR InterPro: IPR006568; PSP.
DR SMART: SM00581; PSP; 1.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51129 MW; 891FCBBB70026D4C CRC64;

```

```

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 469;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 GAVVNC 8
DB 389 GAVVNC 395

```

```

RESULT 11
O9NSS2 PRELIMINARY; PRT; 469 AA.
AC O9NSS2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Well B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL157433; CAB75657.1; -.
DR InterPro: IPR006568; PSP.
DR SMART: SM00581; PSP; 1.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51097 MW; 891FCBBB6CFB8B CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 469;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 GAVVNC 8
DB 389 GAVVNC 395

```

```

RESULT 12
O8N2K5 PRELIMINARY; PRT; 707 AA.
AC O8N2K5;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FLJ90157.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hito Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahashi K.,
RA Masuko Y., Ota T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEBO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK074638; BAC11105.1; -.
DR InterPro: IPR006568; PSP.
DR InterPro: IPR001878; znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00581; PSP; 1.
DR SMART: SM00343; znf_C2HC; 1.
DR PROSITE: PS0158; zf_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 78505 MW; 7BB6E30CF19B2773 CRC64;

```

```

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 707;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 GAVVNC 8
DB 627 GAVVNC 633

```

RESULT 13

096SK7 PRELIMINARY: PRT: 707 AA.

ID 096SK7
AC 096SK7
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FLJ14796.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Isegaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K., Tanai H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuko Y., Kanehori K.;
RT "MEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL AK027702; BAB55308.1;
DR InterPro: IPR006568; PSP.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; zfc_CCHC; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00581; PSP; 1.
DR SMART: SM00343; Znf_C2HC; 1.
DR PROSITE: PS50158; ZF_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 78475 MW; ED4E954E42672F56 CRC64;

Query Match

Best Local Similarity 75.5%; Score 37; DB 4; Length 707;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAVVPC 8
Db 627 GSVVPC 633

RESULT 14

ID 09TRY1 PRELIMINARY: PRT: 29 AA.
AC 09TRY1
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Insulin-like growth factor-binding protein-6, IGFBP-6 (Fragment).
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9826;
[1]
RP SEQUENCE.
RX MEDLINE=92049376; PubMed=1719383;
RA Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding protein-6."
RL Mol. Endocrinol. 5:938-948(1991).
DR HSSP: P24593; IBOE.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR Pfam: PF00219; IGFBP; 1.
FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3129 MW; C976EC5ADE051F6 CRC64;

Query Match

Best Local Similarity 73.5%; Score 36; DB 6; Length 29;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGAVVPC 8
Db 5 CGVTPPC 12

RESULT 15

ID 08UD78 PRELIMINARY: PRT: 61 AA.
AC 08UD78
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein Atu2249.
GN ATU2249 OR AGR C 4091.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium.
NCBI_TaxID=176299;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr., Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C., Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A., Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nestor E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Ouriollo B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L., Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G., Wollan C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B., Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.;
RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL: AE009174; AA143338.1;
DR EMBL: AE008140; AA87991.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 61 AA; 6966 MW; B72DF39E30D5A1E2 CRC64;

Query Match

Best Local Similarity 73.5%; Score 36; DB 16; Length 61;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGAVVPC 8
Db 6 CGTLVPC 13

Search completed: September 4, 2003, 21:06:47
Job time : 34.9144 secs

THIS PAGE BLANK (USPTO)

[illegible]

PI Miltis TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived peptide.
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 39; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVPNC 6
DB 1 CVPNC 6
RESULT 2
AAE16728
ID AAE16728 standard; peptide; 6 AA.
XX
AC AAE16728;
XX
DT 09-APR-2002 (first entry)
XX
DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #8.
XX
KM Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
XX vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KM coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KM hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KM restenosis post-angioplasty; antiaingular; anticoagulant; thrombolytic;
XX cyclic.
XX
OS Ligamentum nuchae.
OS Synthetic.
XX
XX
FH Key
FH Modified-site 1 Location/Qualifiers
FT /note= "This residue forms a cyclic structure with Cys
FT at position 6 with copper as a chelating agent"
FT
FT Modified-site 6 /note= "This residue forms a cyclic structure with Cys
FT at position 1 with copper as a chelating agent"
XX
PN WO200191700-AZ.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17384.
XX
PR 30-MAY-2000; 2000US-0580110.
PR 30-MAY-2000; 2000US-0580156.
PR 30-MAY-2000; 2000US-0580893.

PR 30-MAY-2000; 2000US-0584001.
XX
XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX
XX
PI Miltis TF, Sandberg LB, Jimenez F;
XX
DR WPI: 2002-106259/14.
XX
PT Composition for improving the elasticity of tissue comprises peptide
XX substances which mimic the sequences found in elastin
XX
PS Claim 24; Page 21; 53pp; English.
XX
CC The present invention relates to a composition for the treatment of
CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC their biological equivalents. The composition is useful not only for
CC improving the cosmetic appearance of skin, but for the treatment of
CC various skin disorders associated with reduced elastin formation and for
CC numerous other tissue disorders especially of blood vessels including
CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC angioplasty. They may also be useful in coating surgical appliances i.e.
CC stents. The use of peptides and peptide fragments which mimic those found
CC in elastin overcomes the problem of solubility of the dense cross-linked
CC structure of elastin itself, which renders it difficult to use
CC therapeutically. The peptides are found to improve elastin production in
CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC elastin peptide (HEP), VVPN derived cyclic peptide.
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 39; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVPNC 6
DB 1 CVPNC 6
RESULT 3
AAM83563
ID AAM83563 standard; Protein; 56 AA.
XX
AC AAM83563;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:11156.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis.
XX
XX Homo sapiens.
XX
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226379.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
XX
PI
PI
DR
DR
XX
XX
PT
PT
XX
XX
PS
Claim 11: SEQ ID NO 11156; 3071bp + Sequence Listing; English.
AKK54951 to AKK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to

PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139751.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150386.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156456.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 35; DB 21; Length 342;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6
 1:1111
 DB 159 CMIPNC 164

RESULT 6
AAG18364
ID AAG18364 standard; Protein: 451 AA.
XX
AC AAG18364;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 19739.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152263.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 05-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 12-OCT-1999; 99US-0158232.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 35; DB 21; Length 451;
 Best Local Similarity 66.7%; Pred. No. 3.8e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CVPNC 6
 Db 268 CMIPNC 273

RESULT 7
 AAG18363
 ID AAG18363 standard; Protein: 477 AA.
 XX
 AC AAG18363;
 XX

DT 17-OCT-2000 (First entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 19738.
 DE
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 XX 06-SEP-2000.
 PD
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.

```

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142555.
PR 06-JUL-1999; 99US-0142855.
PR 09-JUL-1999; 99US-0142990.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

```

```

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      89.7%; Score 35; DB 21; Length 477;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPENC 6
Db 294 CMIPNC 299

RESULT 8
ABb62602
ID ABb62602 standard; Protein; 613 AA.
XX
XX ABb62602;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 14598.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX

```


OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB: ABL06705.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 14598: 21pp + Sequence Listing: English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 613 AA:
 QY
 Db 1 CVPNC 6
 1
 1
 1
 1
 3 CAVPNC 8
 RESULT 9
 ABJ04660
 ID ABJ04660 standard: Protein: 394 AA.
 XX
 AC ABJ04660:
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE Protein of NOVX 17 SEQ ID NO 42.
 XX
 XX Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antileptemic;
 KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
 KW tranquiliser; neuroleptic; antidiabetic; antitumor; antiinflammatory;
 KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
 KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
 KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
 KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
 KW epilepsy; stroke; mental disorder; schizophrenia; disorders; golter;
 KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
 KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
 KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
 KW
 XX
 OS Unidentified.
 XX

PN WO200246409-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46586.
 XX
 PR 06-DEC-2000; 2000US-251660P.
 PR 12-DEC-2000; 2000US-255029P.
 PR 08-JAN-2001; 2001US-260326P.
 PR 24-JAN-2001; 2001US-263800P.
 PR 20-FEB-2001; 2001US-269942P.
 PR 24-APR-2001; 2001US-286183P.
 PR 20-AUG-2001; 2001US-313627P.
 PR 12-SEP-2001; 2001US-318712P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Li L, Patturajan M, Shinkets RA, Casman SJ, Malpankar UM;
 PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;
 PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Ganggilli EA;
 PI Boldog FL, Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK;
 PI Zernusen BD;
 XX
 DR WPI: 2002-547774/58.
 DR N-PSDB: APT05473.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
 PT metabolic, neurodegenerative, immune and hematopoietic disorders -
 XX
 PS Claim 1; Page 156; 421pp: English.
 XX
 CC The invention relates to an isolated polypeptide, designated NOVX,
 CC comprising a sequence fully defined in the specification. The isolated
 CC protein, its encoding polynucleotide or an antibody created from the
 CC protein is useful in the manufacture of a medicament for treating a
 CC syndrome associated with a human disease, preferably a NOVX-associated
 CC disorder, or for treating or preventing a NOVX-associated disorder in a
 CC subject, preferably human. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are also useful
 CC for treating or preventing metabolic disorders, diabetes, obesity,
 CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
 CC disease, Parkinson's disorder, immune disorders, haematopoietic
 CC disorders, and various dyslipidaemias, metabolic disturbances associated
 CC with obesity, the metabolic syndrome X, wasting disorders associated with
 CC chronic diseases, and cancer. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are useful for
 CC treating or preventing neurological disorders such as epilepsy, stroke,
 CC mental disorders including mood, anxiety, schizophrenic disorders,
 CC disorders of vesicular transport such as cystic fibrosis, diabetes
 CC mellitus, golter, gastrointestinal disorders including ulcerative
 CC colitis, other conditions associated with abnormal vesicle trafficking
 CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
 CC arthritis. A cell comprising the vector of the invention is useful for
 CC producing non-human transgenic animals. The polynucleotide of the
 CC invention can be used to treat disorders by gene therapy. This sequence
 CC represents one of the isolated NOVX proteins of the invention.
 CC
 XX
 SQ Sequence 394 AA:
 QY
 Db 1 CVPNC 6
 1
 1
 1
 1
 124 CVPNC 129
 RESULT 10
 ABP40414
 ID ABP40414 standard: Protein: 405 AA.
 XX

AC ABP040414;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5259.
 XX
 KW Staphylococcus epidermidis: open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI: 2002-1381255/41.
 DR N-PSDB; ABN92959.
 XX
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -
 XX
 PS Disclosure; SEQ ID 5259; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 SQ Sequence 405 AA;
 XX
 QY 1 CVPVNC 6
 II III
 DB 210 CVCPCNC 215

Query Match 87.2%; Score 34; DB 23; Length 405;
 Best Local Similarity 83.3%; Pred. No. 5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
 AAE16714
 ID AAE16714 standard; peptide; 6 AA.
 XX
 AC AAE16714;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPQ derived peptide #8.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiac;
 KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;
 KW restenosis post-angioplasty; antiauginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.

OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..6
 XX
 PN WO200191700-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17384.
 XX
 PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX
 PI Mlts TF, Sandberg LB, Jimenez F;
 XX
 DR WPI: 2002-106259/14.
 XX
 PT Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 20; 53pp; English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.
 XX
 SQ Sequence 6 AA;
 XX
 QY 1 CVPVNC 6
 IIII I
 DB 1 CVPVPC 6

Query Match 84.6%; Score 33; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
 AAE16717
 ID AAE16717 standard; peptide; 6 AA.
 XX
 AC AAE16717;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPQ derived peptide #11.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiac;
 KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;
 KW restenosis post-angioplasty; antiauginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "This residue forms a cyclic structure with Cys
 FT at position 6 with copper as a chelating agent"
 FT Modified-site 6
 FT /note= "This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 PN WO200191700-A2.
 XX 06-DEC-2001.
 PD 30-MAY-2001; 2001WO-US17384.
 XX 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 XX Mits TF, Sandberg LB, Jimenez F;
 PI WPI; 2002-106259/14.
 DR Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 PT Claim 24; Page 20; 53pp; English.
 PS The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPO derived cyclic peptide.
 CC
 XX Sequence 6 AA:
 SQ
 Query Match 84.6%; Score 33; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CVVPNC 6
 DB 1 CVVPNC 6

PN WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 BA Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS83278.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PT Claim 20; SEQ ID NO 49450; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/publised_pct_sequences.
 CC
 XX Sequence 60 AA:
 SQ
 Query Match 84.6%; Score 33; DB 22; Length 60;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CVVPNC 6
 DB 20 CVVPNC 25

RESULT 13
 ABG19091
 ID ABG19091 standard; Protein: 60 AA.
 XX
 AC ABG19091;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19082.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX

RESULT 14
 AAU42280
 ID AAU42280 standard; Protein: 67 AA.
 XX
 AC AAU42280;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #3176.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX

PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI WPI: 2001-616774/71.
 DR N-PSDB; AAS59516.
 DR
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 34/75; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 67 AA;
 Query Match 84.6%; Score 33; DB 22; Length 67;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CVPVNC 6
 DB 2 CVPVGC 7
 II III I
 RESULT 15
 ID ABB69636 standard; Protein; 106 AA.
 XX ABB69636;
 AC
 XX ABB69636;
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1683.
 XX
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;

KW antiarthritic.
 XX
 OS Homo sapiens.
 XX
 PN WO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002WO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YF, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 DR WPI: 2002-759812/82.
 DR N-PSDB; AB211853.
 DR
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 9; SEQ ID NO 1683; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (AB21119-AB212066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 106 AA;
 Query Match 84.6%; Score 33; DB 23; Length 106;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CVPVNC 6
 DB 72 CVPVNC 77
 II III I

Search completed: September 4, 2003, 21:00:35
 Job time : 30.8824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 9.35294 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-63

Perfect score: 39

Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	97.4	135	2	JC2211
2	38	97.4	135	2	JC2209
3	36	92.3	303	2	DB3082
4	36	92.3	442	1	UKPG
5	35	89.7	334	2	A48151
6	35	89.7	472	2	T47436
7	35	89.7	505	2	C82216
8	35	89.7	507	2	A83105
9	34	87.2	712	2	T16338
10	33	84.6	106	1	REHU36
11	33	84.6	106	1	REHU36
12	33	84.6	365	2	F70988
13	33	84.6	418	2	F70940
14	33	84.6	442	2	T37733
15	33	84.6	782	2	A61625
16	33	84.6	819	2	DB5440
17	33	84.6	905	2	T38980
18	33	84.6	952	2	T03158
19	33	84.6	2180	2	A47651
20	32	82.1	295	2	T22039
21	32	82.1	328	2	B65086
22	32	82.1	328	2	A85959
23	32	82.1	328	2	A91114
24	32	82.1	328	2	A90885
25	32	82.1	407	2	T51226
26	32	82.1	436	2	H87793
27	32	82.1	439	2	T43813
28	32	82.1	491	2	R87793
29	32	82.1	497	2	G87793

30	32	82.1	574	1	S76132	protein kinase pkn
31	32	82.1	1476	2	A45773	kelch protein, lon
32	32	82.1	1912	2	T29088	vitellogenin 1 pre
33	31	79.5	118	2	I35515	dynein-like protei
34	31	79.5	151	2	C69179	hypothetical prote
35	31	79.5	339	2	S73485	heterocyst maturat
36	31	79.5	530	2	A54965	probable desminase
37	31	79.5	556	2	H69279	adenine vesicular
38	31	79.5	756	2	C91266	biodegradative arg
39	31	79.5	756	2	H86106	biodegradative arg
40	31	79.5	756	2	S56346	arginine decarboxy
41	31	79.5	770	1	TMBY22	transcription fact
42	31	79.5	1136	2	T30878	dynein heavy chain
43	31	79.5	1498	2	E86302	hypothetical prote
44	30	76.9	103	2	S70182	hypothetical prote
45	30	76.9	123	2	T01750	glucose-1-phosphat

ALIGNMENTS

RESULT 1	JC2211	hypothetical 14.8K protein, LIM5 - trumpet lily (fragment)
C:Species:	Lilium longiflorum (trumpet lily)	
C:Date:	16-Jul-1999	#sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession:	JC2211	
R:Kobayashi, T.; Kobayashi, E.; Sato, S.; Hotta, Y.; Miyajima, N.; Tanaka, A.; Tabata	DNA Res. 1, 15-26, 1994	
A:Title:	Characterization of cDNAs induced in meiotic prophase in lily microsporocyte	
A:Reference number:	PC2136; PMID:96051386; PMID:7584025	
A:Accession:	JC2211	
A:Status:	preliminary	
A:Molecule type:	mRNA	
A:Residues:	1-135 <KOB>	
A:Cross-references:	DBJ:D21811; NID:g431151; PIDN:BA04835.1; PID:g431152	
Query Match	97.4%	Score 38; DB 2; Length 135;
Best Local Similarity	83.3%	Pred. No. 4.3;
Matches	5; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
OY	1 CVPNC 6	
DB	130 CVPNC 135	
RESULT 2	JC2209	hypothetical 14.9K protein, LIM3 - trumpet lily (fragment)
C:Species:	Lilium longiflorum (trumpet lily)	
C:Date:	16-Jul-1999	#sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession:	JC2209	
R:Kobayashi, T.; Kobayashi, E.; Sato, S.; Hotta, Y.; Miyajima, N.; Tanaka, A.; Tabata	DNA Res. 1, 15-26, 1994	
A:Title:	Characterization of cDNAs induced in meiotic prophase in lily microsporocyte	
A:Reference number:	PC2136; PMID:96051386; PMID:7584025	
A:Accession:	JC2209	
A:Status:	preliminary	
A:Molecule type:	mRNA	
A:Residues:	1-135 <KOB>	
A:Cross-references:	GB:D21819; NID:g442467; PIDN:BA04843.1; PID:g452595	
Query Match	97.4%	Score 38; DB 2; Length 135;
Best Local Similarity	83.3%	Pred. No. 4.3;
Matches	5; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
OY	1 CVPNC 6	
DB	130 CVPNC 135	
RESULT 3	DB3082	

probable permease of ABC transporter PA4504 [imported] - *Pseudomonas aeruginosa* (strain C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83082
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: D83082
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-303 <STO>
 A:Cross-references: GB:AE004864; GB:AE004091; NID:g9950740; PIDN:AA07892.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA4504
 C:Superfamily: oligopeptide permease protein oppB

Query Match 92.3%; Score 36; DB 2; Length 303;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
 DB 211 CVPNC 216

RESULT 4

u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: *Sus scrofa domestica* (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MUID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NAG1>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241
 C:Genetics:
 A:introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringe; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringe homology <RG>
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:179-310-220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
 F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 92.3%; Score 36; DB 1; Length 442;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
 DB 148 CVPNC 153

RESULT 5

A48151

sperm tail protein Mst98Ca - fruit fly (*Drosophila melanogaster*)
 C:Species: *Drosophila melanogaster*
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
 C:Accession: A48151; S23903
 R:Schafer, M.; Borsch, D.; Hulster, A.; Schafer, U.
 Mol. Cell. Biol. 13, 1708-1718, 1993
 A:Title: Expression of a gene duplication encoding conserved sperm tail proteins is t
 A:Reference number: A48151; MUID:93180818; PMID:8441407
 A:Accession: A48151
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <SCH>
 A:Cross-references: EMBL:X67704; NID:98267; PID:98268
 C:Note: sequence extracted from NCBI backbone (NCBIP:125984)
 C:Genetics:
 A:Gene: Mst98Ca
 A:Cross-references: FlyBase:FBgn0002865

Query Match 89.7%; Score 35; DB 2; Length 334;
 Best Local Similarity 83.3%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 DB 242 CVPNC 247

RESULT 6

protein kinase-like protein - *Arabidopsis thaliana*
 N:Alternate names: protein T18B22.10
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
 C:Accession: T47436
 R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke,
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24467
 A:Accession: T47436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-472 <JOP>
 A:Cross-references: EMBL:AL138652
 A:Experimental source: cultivar Columbia; BAC clone T18B22
 C:Genetics:
 A:Map position: 3
 A:introns: 182/2; 327/1
 A:Note: T18B22.10
 C:Superfamily: probable serine/threonine-specific protein kinase ATPK64; protein kina

Query Match 89.7%; Score 35; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 44;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
 DB 268 CVPNC 273

RESULT 7

probable fumarate hydratase, class I Vcl1304 [imported] - *Vibrio cholerae* (strain N169
 C:Species: *Vibrio cholerae*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: C82216
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: AB2035; MUID:20406833; PMID:10952301
 A:Accession: C82216
 A:Status: preliminary
 A:Molecule type: DNA

RESULT 5
 A48151

A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 272 CVPNC 277

RESULT 8
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 272 CVPNC 277

RESULT 9
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6

Db 425 CVPNC 430

RESULT 10
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 272 CVPNC 277

RESULT 11
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 72 CVPNC 77

RESULT 12
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

Query Match 89.7%; Score 35; DB 2; Length 505;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6
Db 72 CVPNC 77

RESULT 13
A:Residues: 1-505 <HEI>
A:Cross-references: GB:AE004210; GB:AE003852; NID:g9655789; PIDN:AAF94463.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1304
A:Map position: 1
C:Superfamily: Iron-dependent fumarate hydratase; Iron-dependent tartarate dehydratase a1

C:Accession: F70988
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A: Reference number: A70500; M0ID: 98295987; PMID: 9634230
 A: Accession: F70988
 A: Status: preliminary; nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1365 <COL>
 A: Cross-references: GB: D95890; GB: AL123456; NID: g3242245; PIDN: CAB09341.1; PID: e1300741;
 A: Experimental source: strain H37Rv
 C: Genetics:
 A: Gene: RV1765c
 C: Superfamily: Mycobacterium tuberculosis hypothetical protein RV1765c

Query Match 84.6%; Score 33; DB 2; Length 365;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 309 CVPNC 314

RESULT 13

F70940
 hypothetical protein RV2015c - Mycobacterium tuberculosis (strain H37Rv)

C: Species: Mycobacterium tuberculosis
 C: Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 15-Sep-2000
 C: Accession: F70940
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A: Reference number: A70500; M0ID: 98295987; PMID: 9634230
 A: Accession: F70940
 A: Status: preliminary; nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1418 <COL>
 A: Cross-references: GB: AL021899; GB: AL123456; NID: g3242282; PIDN: CA11229.1; PID: e125199
 A: Experimental source: strain H37Rv
 C: Genetics:
 A: Gene: RV2015c
 C: Superfamily: Mycobacterium tuberculosis hypothetical protein RV1765c

Query Match 84.6%; Score 33; DB 2; Length 418;
 Best Local Similarity 83.3%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 309 CVPNC 314

RESULT 14

T37733
 hypothetical zinc finger protein - fission yeast (Schizosaccharomyces pombe)

C: Species: Schizosaccharomyces pombe
 C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C: Accession: T37733
 R: Davis, P.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A: Reference number: Z21741
 A: Accession: T37733
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1442 <RAV>
 A: Cross-references: EMBL: AL121745; PIDN: CAB57403.1; GSPDB: GNO00066; SPDB: SPAC16.05c

A: Experimental source: strain 972h-; cosmid c16
 C: Genetics:
 A: Gene: SPDB: SPAC16.05c
 A: Map position: 1

Query Match 84.6%; Score 33; DB 2; Length 442;
 Best Local Similarity 83.3%; Pred. No. 97;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 352 CVPNC 357

RESULT 15

A61625
 tenascin-like protein precursor - fruit fly (Drosophila melanogaster)

C: Species: Drosophila melanogaster
 C: Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 11-Jan-2000
 C: Accession: A61625; S28463
 R: Baumgartner, S.; Chiquet-Ehrismann, R. Mech. Dev. 40, 165-176, 1993
 A: Title: Ten(a), a Drosophila gene related to tenascin, shows selective transcript 10
 A: Reference number: A61625; M0ID: 93264270; PMID: 7684246
 A: Accession: A61625
 A: Status: preliminary
 A: Molecule type: mRNA
 A: Residues: 1782 <BAU>
 A: Cross-references: EMBL: X68794
 C: Genetics:
 A: Gene: ten-a
 A: Cross-references: FlyBase: FBgn000446
 C: Superfamily: unassigned EGF-related proteins; EGF homology
 C: Keywords: tandem repeat
 F: 1-62/Domain: signal sequence #status predicted <Sig>
 F: 63-782/Product: tenascin-like protein #status predicted <MAT>
 F: 497-524/Domain: EGF homology <EGF>

Query Match 84.6%; Score 33; DB 2; Length 782;
 Best Local Similarity 83.3%; Pred. No. 16e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6
 Db 497 CVPNC 502

Search completed: September 4, 2003, 21:10:53
 Job time: 10.3529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 ; Search time 5.20588 Seconds
(Without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893d-63
Perfect score: 39
Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	92.3	442	1	UROK_PIG
2	33	84.6	105	1	RL44_HUMAN
3	33	84.6	819	1	AHML_ARATH
4	33	84.6	905	1	YAG1_SCHPO
5	33	84.6	1969	1	2292_HUMAN
6	32	82.1	328	1	HYBA_ECOLI
7	32	82.1	439	1	IF2G_ENCCU
8	32	82.1	574	1	SPKB_SYNY3
9	32	82.1	1477	1	KELC_DROME
10	32	82.1	1912	1	VITI_CHICK
11	31	79.5	339	1	Y467_MYCPN
12	31	79.5	344	1	ZDH4_HUMAN
13	31	79.5	423	1	WR58_ARATH
14	31	79.5	506	1	GAE_HUMAN
15	31	79.5	556	1	ADEC_ARCFU
16	31	79.5	652	1	P52K_MOUSE
17	31	79.5	755	1	ADIA_ECOLI
18	31	79.5	761	1	P52K_HUMAN
19	31	79.5	770	1	ACE2_YEAST
20	31	79.5	8545	1	ANCI_DROME
21	30	76.9	125	1	GLGS_MAIZE
22	30	76.9	146	1	VA28_VACCV
23	30	76.9	277	1	UBS1_YEAST
24	30	76.9	309	1	WR26_ARATH
25	30	76.9	311	1	Y467_MYCCE
26	30	76.9	332	1	PTA_CLOTH
27	30	76.9	398	1	ACTU_DROME
28	30	76.9	403	1	ARGD_BROME
29	30	76.9	451	1	Y996_SYNY3
30	30	76.9	463	1	D5DR_FUGRU
31	30	76.9	473	1	GLGS_WHEAT
32	30	76.9	483	1	GLGS_ORISA
33	30	76.9	489	1	GLGS_BETVU

34	30	76.9	506	1	GAE_RAT	Q9es14 ratius norv
35	30	76.9	508	1	GLGS_VICFA	P52416 vicla faba
36	30	76.9	512	1	GLGT_VICFA	P52417 vicla faba
37	30	76.9	512	1	WR33_ARATH	O88905 arabidopsis
38	30	76.9	513	1	GLGS_HORVU	P55238 hordeum vul
39	30	76.9	520	1	GLGS_ARATH	P55228 arabidopsis
40	30	76.9	520	1	GLGS_BRANA	Q9m462 brassica na
41	30	76.9	521	1	GLGS_LYCES	Q42882 lycopersico
42	30	76.9	521	1	GLGS_SOLITU	P23509 solanum tub
43	30	76.9	525	1	ACUB_NECR	P15937 neurospora
44	30	76.9	606	1	SP2_HUMAN	Q02086 homo sapien
45	30	76.9	644	1	BTD_DROME	Q24266 drosophila

ALIGNMENTS

RESULT 1	UROK_PIG	STANDARD:	PRT:	442 AA.
ID	UROK_PIG			
AC	P04185:			
DT	20-MAR-1987 (Rel. 04, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)			
GN	(U-plasminogen activator).			
GN	PLAU.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=85087954; PubMed=6096832;			
RA	Nagamine Y., Pearson D., Altus M.S., Reich E.;			
RT	"cDNA and gene nucleotide sequence of porcine plasminogen activator.";			
RL	Nucleic Acids Res. 12:9525-9541(1984).			
RN	[2]			
RP	REVISION TO 241.			
RA	Nagamine Y.;			
RL	Submitted (DEC-1986) to the PIR data bank.			
CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in			
CC	plasminogen to form plasmin.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-1- SIMILARITY: Contains 1 EGF-like domain.			
CC	-1- SIMILARITY: Contains 1 EGF-like domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X01648; CAA25806.1; -			
DR	EMBL: X02724; CAA26511.1; -			
DR	PIR: A00932; URPG.			
DR	HSSP: P00749; IKDU.			
DR	MEROPS: S01.231; -			
DR	InterPro: IPR001314; Chymotrypsin.			
DR	InterPro: IPR006209; EGF like.			
DR	InterPro: IPR000001; Kringle.			
DR	InterPro: IPR001254; Ser-protease-Try.			
DR	Pfam: PF00051; Kringle; 1.			
DR	Pfam: PF00089; trypsin; 1.			
DR	PRINTS: PR00722; CHYMOTRYPSIN.			
DR	PRINTS: PR00018; KRINGLE.			
DR	PRODOM: PD000395; Kringle; 1.			
DR	SMART: SM00130; KR; 1.			
DR	SMART: SM00020; TRYD-Spc; 1.			
DR	PROSITE: PS00022; EGF_1; 1.			

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE L44E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15661; AA36589.1; -
 DR EMBL: U78027; AAB64204.1; -
 DR EMBL: BC001781; AA01781.1; -
 DR EMBL: U58105; AAB47245.1; -
 DR EMBL: AK002540; BAB22175.1; -
 DR EMBL: AK003166; BAB22616.1; -
 DR EMBL: AK010629; BAB27075.1; -
 DR EMBL: AK012502; BAB28285.1; -
 DR EMBL: BC019810; AAH19810.1; -
 DR EMBL: BC027515; AAH27515.1; -
 DR EMBL: M19635; AAB54277.1; -
 DR EMBL: AB000910; BAA19210.1; -
 DR PIR: A29820; R6RT36.
 DR Genew: HGNC:10359; RPL36A.
 DR GK: P09896; -
 DR MIM: 180469; -
 DR MGD: MGI:1201789; RPL44.
 DR GO: GO:0005840; C:Ribosome; NAS.
 DR InterPro: IPR000552; Ribosomal_L44E.
 DR Pfam: PF00935; Ribosomal_L44; 1.
 DR ProDom: PD002841; Ribosomal_L44E; 1.
 DR PROSITE: PS01172; RIBOSOMAL_L44E; 1.
 DR KW Ribosomal protein.
 FT INIT_MET 0
 FT CONFLICT 37
 FT SEQUENCE 105 AA; 12310 MW; 4C599DD5924C4EB3 CRC64;
 SQ
 Query Match 84.6%; Score 33; DB 1; Length 105;
 Best Local Similarity 83.3%; Pred. NO. 5;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 CVPNMC 6
 Db 71 CVPNMC 76
 RESULT 3
 AHM1_ARATH STANDARD; PRT; 819 AA.
 ID AHM1_ARATH STANDARD; PRT; 819 AA.
 AC O9M3H5; O9SM66;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-SEP-2003 (Rel. 42; Last annotation update)
 DE Potential cadmium/zinc-transporting ATPase HMA1 (EC 3.6.3.3)
 DE (EC 3.6.3.5)
 GN HMA1 OR AT4G37270 OR C7A10.90 OR AP22.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 CC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=2008348; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Voicaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grievell L.A., Rieger M.,
 RA Weichselgattner M., de Simone V., Obermaier B., Macho R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
 RA Lantham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braken M., Melijens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weichselgatter T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirse W.,
 RA Moeljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernieris S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McKay K., Mayes R.,
 RA Petrelli A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
 RA Borikova D., Bloecker H., Scharfe M., Grimm M., Loehner T.H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fairman B., Grandeth K., Dauner D., Herzl A.,
 RA Neumann S., Argitrou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Mendenhall A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedford F., Cooke R., Berger C., Montfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse C.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bleike C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller R., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vall D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Holtman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochl M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT *Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana*;
 RL Nature 402:769-777(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RA Page S.L., Pittman J.K., Krilger G.C., Williams L.E.;
 RT *Identification of a putative heavy metal P-type ATPase in
 Arabidopsis*;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN CADMIUM/ZINC TRANSPORT (POTENTIAL).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Cd(2+)(In) -> ADP + phosphate +
 Cd(2+)(Out).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Zn(2+)(In) -> ADP + phosphate +
 Zn(2+)(Out).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 ATPases). Subfamily IB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z99707; CAB16773.1; -
 DR EMBL: AL161591; CAB80393.1; -
 DR EMBL: AJ400906; CAB90352.1; -
 DR PIR: D85440; D85440.
 DR InterPro: IPR005416; ATPase-IB_hvy.
 DR InterPro: IPR001475; ATPase-ET-2.
 DR InterPro: IPR006404; Heavy_mel_ATPase.
 DR InterPro: IPR005834; Hydrolase.
 DR Pfam: PF00122; Et-E2_ATPase; 1.

DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATAPASE.
 DR TIGRGRAMS: TIGR01512; ATPase-IB2.Cd; 1.
 DR TIGRGRAMS: TIGR01535; ATPase-IB_hvy; 1.
 DR TIGRGRAMS: TIGR01494; ATPase-P-type; 2.
 DR PROSITE: PS00154; ATPase_EI_E2; 1.
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding; Magnesium;
 ZINC: Cadmium; Metal-binding; Multigene family.
 FT DOMAIN 1 122 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 123 144.. EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 145 153 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 154 173 POTENTIAL.
 FT DOMAIN 174 180 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 181 201 POTENTIAL.
 FT DOMAIN 202 202 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 203 223 POTENTIAL.
 FT TRANSMEM 224 361 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 362 384 POTENTIAL.
 FT DOMAIN 385 398 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 399 416 POTENTIAL.
 FT TRANSMEM 417 737 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 738 757 POTENTIAL.
 FT DOMAIN 758 762 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 763 781 POTENTIAL.
 FT DOMAIN 782 819 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 79 89 POLY-HIS.
 FT MOD_RES 809 817 PHOSPHORYLATION (BY SIMILARITY).
 FT METAL 682 682 MAGNESIUM (BY SIMILARITY).
 FT METAL 686 686 MAGNESIUM (BY SIMILARITY).
 FT VARIANT 53 53 N -> S (IN STRAIN LANDSBERG ERECTA).
 FT VARIANT 105 105 M -> V (IN STRAIN LANDSBERG ERECTA).
 FT VARIANT 659 659 P -> S (IN STRAIN LANDSBERG ERECTA).
 SQ SEQUENCE 819 AA; 88186 MW; F281EABF33C0E52 CRC64;

Query Match 84.6%; Score 33; DB 1; Length 819;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CYPNC 6
 Db 485 CCIPNC 490

RESULT 4

YAGI_SCHPO STANDARD; PRT; 905 AA.
 ID YAGI_SCHPO
 AC 009866; 090UH9;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical WD-repeat protein C12G12.0lc in chromosome I.
 GN SPAC12G12.0lc OR SPAC630.02
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; Pubmed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Soudras J., Peat N., Hayes S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer C., Mostl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambolt R., Furrer B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Contains 4 WD repeats.
 CC -1- SIMILARITY: TO YEAST YBL104C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL109832; CAB52724.1;
 DR EMBL: 265568; CAA91496.1;
 DR PIR: T38980; T38980
 DR GeneDB:SPDome; SPAC12G12.0lc;
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF004400; WD40;
 DR SMART: SM00320; WD40; 4.
 DR PROSITE: PS00678; WD_REPEATS.1; FALSE_NEG.
 DR PROSITE: PS50082; WD_REPEATS.2; FALSE_NEG.
 DR PROSITE: PS50294; WD_REPEATS_REGION; FALSE_NEG.
 KW Hypothetical protein; Repeat; WD repeat.
 FT REPEAT 42 82 WD 1.
 FT REPEAT 86 128 WD 2.
 FT REPEAT 136 175 WD 3.
 FT REPEAT 177 217 WD 4.
 SQ SEQUENCE 905 AA; 101948 MW; 56BCF35CA07F709B CRC64;

Query Match 84.6%; Score 33; DB 1; Length 905;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CYPNC 6
 Db 893 CYPNC 898

RESULT 5

Z292_HUMAN STANDARD; PRT; 1969 AA.
 ID Z292_HUMAN
 AC 060281; Q9H8G3; Q9H8J4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 292.
 GN ZNF292 OR KIA0530.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE OF 1-874 AND 1160-1786 FROM N.A.
 RC TISSUE=Placenta;
 RC Iso9a T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe M., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuko Y.,
 RA Niinomya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project."; www.genbank.org
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2].
 RP SEQUENCE OF 407-1969 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AK023626; BAB14622.1; -
 DR EMBL: AK023712; BAB14654.1; -
 DR EMBL: AB011102; BAA25456.1; -
 DR HSSP: P08153; INCS.
 DR Genew: HGNC:18410; ZNF292.
 DR InterPro: IPR007087; Znf.C2H2.
 DR Pfam: PF00096; zf-C2H2; 17.
 DR SMART: SM00355; ZNF_C2H2; 12.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 9.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 7.
 KW DNA-binding; regulation; zinc-finger; Metal-binding; Nuclear protein;
 KM DNA-binding; Repeat.
 FT ZN_FING 25 C2H2-TYPE.
 FT ZN_FING 53 77 C2H2-TYPE.
 FT ZN_FING 344 369 C2H2-TYPE.
 FT ZN_FING 1193 1219 C2H2-TYPE.
 FT ZN_FING 1360 1385 C2H2-TYPE.
 FT ZN_FING 1418 1443 C2H2-TYPE.
 FT ZN_FING 1462 1487 C2H2-TYPE.
 FT ZN_FING 1502 1527 C2H2-TYPE.
 FT ZN_FING 1632 1656 C2H2-TYPE.
 FT ZN_FING 601 601 E -> K (IN REF. 1).
 FT CONFLICT 776 776 N -> S (IN REF. 1).
 FT CONFLICT 1291 1291 I -> V (IN REF. 1).
 FT CONFLICT 1408 1408 E -> K (IN REF. 1).
 FT SEQUENCE 1969 AA; 218445 MW; 3D5D25D5ADCFDPA5A CRC64;
 SQ
 Query Match 84.6%; Score 33; DB 1; Length 1969;
 Best Local Similarity 83.3%; Pred. No. 86;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPN6 6
 DB 1195 CVPN6 1200
 ID HYBA_ECOLI STANDARD; PRT; 328 AA.
 AC P31179;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hydrogenase-2 operon protein hybA precursor.
 GN HYBA OR B2996 OR Z4350 OR ECS3881 OR SF3043.
 OS Escherichia coli,

OS Escherichia coli O157:H7, and
 OS Shigella flexneri.
 OC Bacteria: Proteobacteria: Gammaproteobacteria: Enterobacteriales;
 OC Enterobacteriaceae: Escherichia.
 ON NCBI_TaxID=562, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / TGI;
 RX MEDLINE=94292472; PubMed=8021226;
 RA Menon N.K., Chatelet C.Y., Deravantian M., Wendt J.C.,
 RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.;
 RT "Cloning, sequencing, and mutational analysis of the hyb operon
 RT encoding Escherichia coli hydrogenase 2.";
 RL J. Bacteriol. 176:4416-4423(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfale G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao Y., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.H.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 CC -1- FUNCTION: PARTICIPATES IN THE PERIPLASMIC ELECTRON-TRANSFERRING
 CC ACTIVITY OF HYDROGENASE 2 DURING ITS CATALYTIC TURNOVER.
 CC -1- COFACTOR: BINDS 3 4FE-4S CLUSTERS AND A 3FE-4S CLUSTER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE THE SMALL SUBUNIT
 CC OF HYDROGENASE 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U09177; AAA21589.1; -
 DR EMBL: U28377; AAA69163.1; -
 DR EMBL: AE000382; AAC76032.1; -
 DR EMBL: AE005529; AAG38133.1; -
 DR EMBL: AP002563; BAB37304.1; -
 DR EMBL: AE015316; AAM4521.1; -
 DR PIR: A9114; A9114.
 DR PIR: B65086; B65086.
 DR Ecocyc: B611799; hyba.
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR006311; Tat.
 DR Pfam: PF00037; fer4; 1.
 DR TIGRPFAM: TIGR01409; Tat_signal_seq; 1.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; 1.
 KM Oxidoreductase; Signal; Periplasmic; Iron-sulfur; 4Fe-4S; 3Fe-4S;
 KM Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 328
 FT METAL 47 47
 FT METAL 50 50
 FT METAL 53 53
 FT METAL 57 57
 FT METAL 112 112
 FT METAL 115 115
 FT METAL 120 120
 FT METAL 124 124
 FT METAL 145 145
 FT METAL 148 148
 FT METAL 151 151
 FT METAL 155 155
 FT METAL 174 174
 FT METAL 177 177
 FT METAL 193 193
 FT METAL 197 197
 SQ SEQUENCE 328 AA; 36003 MW; 77203A0F50F61662 CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 1; Length 328;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
 11111
 DB 115 CVPNC 120

RESULT 7
 IF2G_ENCCU STANDARD; PRT; 439 AA.
 AC 096719;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Eukaryotic translation initiation factor 2 gamma subunit (eif-2-gamma).
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_Taxid=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=mouse;
 RA Duffieux F., Peyret P., Roe B.A., Vlvares C.P.;
 RT "Putative elongation factor 2, from Encephalitozoon cuniculi.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
 CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
 CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
 CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
 CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
 CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
 CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER

CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
 CC BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
 CC -I- SUBUNIT: Heterotrimer composed of an alpha, a beta, and a gamma
 CC chain.
 CC -I- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EIF2G SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ006823; CAA07260.1; -
 DR PIR: T43813; T43813.
 DR InterPro: IPR000795; EF-GTPbind.
 DR InterPro: IPR004161; EFTU_D2.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR PRINTS: PR00315; ELONGATNFCT.
 KM Initiation factor; Protein biosynthesis; GTP-binding.
 FT NP_BIND 20 27
 FT NP_BIND 103 107
 FT NP_BIND 159 162
 FT NP_BIND 162 162
 SQ SEQUENCE 439 AA; 47846 MW; 339B89D4FD8DFC8B CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 1; Length 439;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
 11111
 DB 67 CVPNC 72

RESULT 8
 SPKB_SYNY3 STANDARD; PRT; 574 AA.
 ID AC P74297;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine-protein kinase B (EC 2.7.1.37).
 GN SPKB OR SLR1697.
 OS Synchocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 OX NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamel A., Ikeuchi M.;
 RT "A novel gene, SPKB, encodes active Ser/Thr protein kinase in the
 RT motile cyanobacterium Synchocystis sp. PCC 6803.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: This protein is a serine/threonine kinase.
 CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -I- SIMILARITY: Contains 2 pentapeptide repeat domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB046598; BAB17034.1; -
 DR EMBL: D90914; BAA18391.1; -
 DR PIR: S76132; S76132.
 DR InterPro: IPR001646; 5peptide.repeat.
 DR InterPro: IPR000719; Set_ltr-kinase.
 DR InterPro: IPR002290; Set_ltr-kinase.
 DR Pfam: PF00805; Pentapeptide; 2.
 DR Pfam: PF00069; Kinase; 1.
 DR PRODOM: PD000001; Prot_kinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW Complete proteome.
 FT DOMAIN 34 301 PROTEIN KINASE.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 65 65 ATP (BY SIMILARITY).
 FT ACT_SITE 163 163 BY SIMILARITY.
 FT DOMAIN 454 493 PENTAPEPTIDE 1.
 FT DOMAIN 504 543 PENTAPEPTIDE 2.
 SO SEQUENCE 574 AA; 63079 MW; D96B2C5ABEE405A3 CRC64;

Query March 82.1%; Score 32; DB 1; Length 574;
 Best Local Similarity 83.3%; Pred. NO. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVPNC 6
 1111
 Db 4 CVPNC 9

RESULT 9
 KEGG DROME STANDARD: PRT; 1477 AA.
 ID KEGG DROME 004653; OGVTA2.
 AC 004653; 004653; OGVTA2.
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ring canal kelch protein [contains: Kelch short protein].
 GN KLC OR CG7210.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=93201592; PubMed=8453663;
 RA Adams M.D., Celitker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris R.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland F.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson S.D., Nelson K.A., Nixon K., Nuskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP CHARACTERIZATION.
 RC TISSUE=Embryo;
 RX MEDLINE=97236487; PubMed=9118811;
 RA Robinson D.N., Cooley L.;
 RT "Examination of the function of two kelch proteins generated by stop
 codon suppression";
 RL Development 124:1405-1417(1997).
 CC -1- FUNCTION: COMPONENT OF RING CANALS THAT REGULATES THE FLOW OF
 CYTOPLASM BETWEEN CELLS. MAY BE INVOLVED IN THE REGULATION OF
 CYTOPLASM FLOW FROM NURSE CELLS TO THE OOCYTE DURING OOGENESIS.
 CC BINDS ACTIN.
 CC -1- SUBCELLULAR LOCATION: INNER SURFACE OF CYTOPLASMIC BRIGES OR RING
 CANALS PRESENT IN EGG CHAMBERS. SUBCORTICALLY IN IMAGINAL DISK
 EPITHELIA.
 CC -1- TISSUE SPECIFICITY: BOTH PROTEINS ARE EXPRESSED IN OVARIES, MALE
 TESTIS, OVARECTOMIZED FEMALES, CUTICLE, SALIVARY GLAND AND
 IMAGINAL DISKS. KELCH ORF1 IS THE PREDOMINANT PROTEIN AND IS ALSO
 EXPRESSED IN FAT BODIES ON ENTRY INTO METAMORPHOSIS LEVELS OF
 FULL LENGTH PROTEIN INCREASE IN TESTIS AND IMAGINAL DISKS.
 CC -1- DEVELOPMENTAL STAGE: LARVAE, PUPAE AND ADULTS.
 CC -1- SIMILARITY: Contains 6 kelch repeats.
 CC -1- CAUTION: REF.3 BELIEVES RESIDUE 690 IS A SELENOCYSTEINE.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L08483; AAA53471.1; -
 CC EMBL: L08483; AAA53472.2; -
 CC EMBL: AE003657; AAF53651.1; ALT_SEQ.
 CC HSSP: Q05516; ICS3.
 CC FlyBase: FBgn0001301; kel.
 CC GO: GO:0007292; P: oogenesis; IMP.
 CC GO: GO:0007301; P: ring canal formation; IDA.
 CC InterPro: IPR000210; BTB_PoZ.
 CC InterPro: IPR006531; Kelch.
 CC InterPro: IPR006552; Kelch_rep.
 CC Pfam: PF00651; BTB; 1.

```

DR Pfam: PF01344; Kelch; 6.
DR PRINTS; PR00501; KELLCHREPEAT.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00612; Kelch; 6.
DR PROSITE; PS0097; BTB; 1.
KM Cytoskeleton; Actin-binding; Selenium; Selenocysteine; Kelch repeat;
KW Repeat.
FT CHAIN 1 1477 KELLCH PROTEIN.
FT CHAIN 1 689 KELLCH SHORT PROTEIN.
FT DOMAIN 157 223 BTB.
FT REPEAT 404 449 KELLCH 1.
FT REPEAT 450 496 KELLCH 2.
FT REPEAT 498 543 KELLCH 3.
FT REPEAT 545 592 KELLCH 4.
FT REPEAT 594 639 KELLCH 5.
FT REPEAT 641 687 KELLCH 6.
FT DOMAIN 18 28 ASN-RICH.
FT DOMAIN 29 87 GLN-RICH.
FT DOMAIN 29 87 POLY-GLN.
FT DOMAIN 78 83 POLY-GLN.
FT SE_CYS 690 690 PROBABLE.
FT CONFLICT 493 493 V -> A (IN REF. 1).
FT CONFLICT 596 596 A -> R (IN REF. 1).
FT CONFLICT 824 824 P -> L (IN REF. 1).
FT CONFLICT 858 858 G -> D (IN REF. 1).
FT CONFLICT 1083 1083 A -> R (IN REF. 1).
FT CONFLICT 1086 1086 A -> G (IN REF. 1).
SQ SEQUENCE 1477 AA; 160086 MW; 4851EEAE9D9DBA47 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 1477;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 1025 CVPNC 1030

RESULT 10
ID VIT1-CHICK STANDARD; PRT; 1912 AA.
AC P87498;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vitellogenin I precursor (Minor vitellogenin) [contains: Lipovitelin
DE I (LVI); Phosvitin (PV); Lipovitelin II (LVII); YGP42].
GN VTG1 OR VTG1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Mabuchi N., Yamamura J.-I., Adachi T., Aoki N., Nakamura R.,
RA Matsuda T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1628-1639, AND IDENTIFICATION OF YGP42.
RX MEDLINE=95322425; PubMed=7599159;
RA Yamamura J.-I., Adachi T., Aoki N., Nakajima H., Nakamura R.,
RA Matsuda T.;
RT "Precursor-product relationship between chicken vitellogenin and the
RT yolk proteins: the 40 kDa yolk plasma glycoprotein is derived from
RT the C-terminal cysteine-rich domain of vitellogenin II.";
RL Blochim. Biophys. Acta 1244:384-394(1995).
CC -1- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEIN THAT ARE
CC SOURCES OF NUTRIENTS DURING EARLY DEVELOPMENT OF OVIPAROUS
CC ORGANISMS.
CC -1- FUNCTION: PHOSVITIN IS BELIEVED TO BE OF IMPORTANCE IN
CC SEQUESTERING CALCIUM, IRON AND OTHER CATIONS FOR THE DEVELOPING

```

```

CC EMBRYO.
CC -1- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD
CC AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENDOCYTOSIS INTO GROWING
CC OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE
CC RESPECTIVE YOLK COMPONENTS.
CC -1- INDUCTION: By steroids (estrogen).
CC -1- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST
CC HIGHLY PHOSPHORYLATED (10%) PROTEINS IN NATURE.
CC -1- PTM: CATHEPSIN D IS RESPONSIBLE FOR INTRAOOCYTIC PROCESSING OF
CC VITELLOGENIN.
CC -1- PTM: MAY CONTAIN INTRAMOLECULAR DISULFIDE BONDS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenselsib-sib.ch).
CC EMBL: D89547; BAA13973.1; -.
CC PIR: T29088; T29088.
DR InterPro: IPR001747; Lipid_transprt_N.
DR InterPro: IPR001846; VWF_D.
DR Pfam: PF01347; Vitellogenin_N; 1.
DR Pfam: PF00094; vwd; 1.
DR SMART: SM00638; LPD_N; 1.
DR SMART: SM00216; VWD; 1.
KM Glycoprotein; Phosphorylation; Storage protein; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 1912 VITELLOGENIN I.
FT CHAIN 16 1139 LIPOVITELIN I (BY SIMILARITY).
FT CHAIN 1140 1391 PHOSVITIN (BY SIMILARITY).
FT CHAIN 1392 1627 LIPOVITELIN II (BY SIMILARITY).
FT CHAIN 1628 1912 YGP42.
FT DOMAIN 1093 1104 POLY-SER.
FT DOMAIN 1150 1289 SER-RICH.
FT DOMAIN 1359 1397 SER-RICH.
FT DOMAIN 1612 1615 POLY-SER.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1206 1206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1375 1375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1662 1662 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1698 1698 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1703 1703 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1912 AA; 210630 MW; 89BA6273D6492E51 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 1912;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 791 CIVPNC 796

RESULT 11
ID Y467_MYCPN STANDARD; PRT; 339 AA.
AC Y467_MYCPN
AC P75110;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein MG467 homolog
DE (K05.orf339).
GN MPN683 OR MP159.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;

```



```

RA MEDLINE-97105885; PubMed-8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000017; AAB95807.1; -
DR PIR; S73485; S73485.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR PRODOM; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT NP_BIND 150 157 ATP (POTENTIAL).
SQ SEQUENCE 339 AA; 38138 MW; DAC0D8F08294E6AB CMC64;

Query Match 79.5%; Score 31; DB 1; Length 339;
Best Local Similarity 83.3%; Pred. NO. 39;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 174 CAVANC 179

RESULT 12
ID ZD4_HUMAN STANDARD; PRT; 344 AA.
AC O9NPG8; O9NPG8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 4 (zinc finger protein 374)
DE (DC1).
GN ZDHHC4 OR ZNF374.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxId=9606;
RX [1]
RP SEQUENCE FROM N.A.
RA Li Y., Peng Y., Li N., Gu W., Han Z., Fu G., Chen Z.;
RT "Novel genes expressed in human dendritic cell."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NMDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166;
Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,

```

```

RA Ansoorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,
RA Wandt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dichtchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carrincci F., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK001341; BAA91636.1; -
DR EMBL; AF201931; AAF86867.1; -
DR EMBL; AL13674; CAB66609.1; -
DR EMBL; BC001239; AAH01239.1; -
DR Genew; HGNC:18471; ZDHHC4.
DR InterPro; IPR001594; Znf-DHHC.
DR Pfam; PF01529; zf-DHHC; 1.
DR PRODOM; PD003041; Znf_DHHC; 1.
DR PROSITE; PS50216; ZF_DHHC; 1.
KW Transmembrane; Zinc-finger.
FT TRANSMEM 3 23
FT TRANSMEM 68 88
FT TRANSMEM 100 120
FT TRANSMEM 193 213
FT TRANSMEM 234 254
FT TRANSMEM 256 276
FT ZN_FING 149 199
FT CONFLICT 244 244 F -> I (IN REF. 3).
FT CONFLICT 276 276 F -> S (IN REF. 3).
FT CONFLICT 294 294 D -> V (IN REF. 3).
SQ SEQUENCE 344 AA; 39786 MW; 302CF22B14036908 CMC64;

Query Match 79.5%; Score 31; DB 1; Length 344;
Best Local Similarity 50.0%; Pred. NO. 40;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
Db 43 CIIPEC 48

```

```

RESULT 13
WR38_ARATH STANDARD: PRT: 423 AA.
AC 03MUT7: 09MAC2:
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 42, Last annotation update)
DE Probable WRKY transcription factor 58 (WRKY DNA-binding protein 58).
GN WRKY58 OR AT3G01080 OR T4P13.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
ON [1]
RX SEQUENCE FROM N.A.
RC STRAIN-CV, Columbia; TISSUE=flower;
RA Kushnir S., Ulker B., Somssich I.E.;
RT "Arabidopsis thaliana transcription factor WRKY58."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, Columbia; PubMed-11130713;
RX MEDLINE-21016720; PubMed-11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unselid M.,
RA Partmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delzeny M., Boutry M., Glivell L.A., Mache R., Pulgomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brotlier P.,
RA Muncker P., Catolico L., Weissenbach J., Saurin W., Queller F.,
RA Wurmacher E., Dzzonek H., Erife H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T., H., Nordstiek G.,
RA Reichelt J., Schafie M., Schoen O., Barques M., Terol J., Clement J.,
RA Navaretto P., Collado C., Perez-Perez A., Oltmawelder B., Duchemin D.,
RA Cooke R., Laurie M., Berger-Llano C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rued S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Uterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltoscher J., Sellers P., Gill J.E., Feldlyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Nakyama C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana."
RL Nature 408:820-822(2000).
CC -1- FUNCTION: Transcription factor. Interacts specifically with the W
box (5'-(T)TGAC(C/T)-3'), a frequently occurring elicitor-
responsive cis-acting element (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 2 WRKY domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: AF426254; AAL29431.1; -
CC DR EMBL: AC008261; AAF26166.1; -
CC DR InterPro: IPR003657; WRKY.
CC DR Pfam: PF03106; WRKY; 2.
CC DR PROSITE: PS50811; WRKY; 2.
CC Transcription regulation; DNA-binding; Nuclear protein; Repeat.
KW

```

```

FT DNA_BIND 161 225 WRKY 1.
FT DNA_BIND 300 365 WRKY 2.
FT DOMAIN 150 154 POLY-ASN.
FT DOMAIN 257 270 POLY-ASP.
FT CONFLICT 81 81 L -> LNTGWSVSPGGGRSTAGMRAGGGMFTIPSGSPS
FT SL (IN REF. 2).
SQ SEQUENCE 423 AA; 47121 MW; EF6C2F1BF3F16B15 CRC64;
Query Match 79.5%; Score 31; DB 1; Length 423;
Best local Similarity 66.7%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CVPNC 6
Db 331 CTPNC 336
RESULT 14
GAE_HUMAN STANDARD: PRT: 506 AA.
AC P78334; O15345; O15346; O99520;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gamma-aminobutyric-acid receptor epsilon subunit precursor (GABA(A)
DE receptor).
GN GABRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
ON [1]
RX SEQUENCE FROM N.A.
RC MEDLINE-97192095; PubMed-9039914;
RA Davies P.A., Hanna M.C., Hales T.G., Kirkness E.F.;
RT "Insensitivity to anaesthetic agents conferred by a class of GABA(A)
RT receptor subunit."
RL Nature 385:820-823(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-97480709; PubMed-9339354;
RA Wilke K., Gaul R., Krausk S.M., Poustka A.;
RT "A gene in human chromosome band Xq28 (GABRE) defines a putative new
RT subunit class of the GABA(A) neurotransmitter receptor."
RL Genomics 45:1-10(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-97238072; PubMed-9084408;
RA Garret M., Basicles L., Boue-Grabot E., Sartor P., Chartron G.,
RA Bloch B., Margolske R.F.;
RT "An mRNA encoding a putative GABA-gated chloride channel is expressed
RT in the human cardiac conduction system."
RL J. Neurochem. 68:1382-1389(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Hanna M.C., Hales T.G., Kirkness E.F.;
RT "Alternative transcripts of a gene encoding the GABA-A receptor
RT epsilon subunit on chromosome Xq28."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GABA, THE MAJOR INHIBITORY NEUROTRANSMITTER IN THE
CC VERTEBRATE BRAIN, MEDIATES NEURONAL INHIBITION BY BINDING TO THE
CC GABA/BENZODIAZEPINE RECEPTOR AND OPENING AN INTEGRAL CHLORIDE
CC CHANNEL.
CC -1- SUBUNIT: GENERALLY PENTAMERIC. ASSOCIATES WITH ALPHA AND BETA
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND IN MANY TISSUE. HIGHEST LEVELS OF
CC EXPRESSION IN ADULT HEART AND PLACENTA.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC -----
DR EMBL: U66661; AAB49284.1; -
DR EMBL: Y09763; CAA70903.1; JOINED.
DR EMBL: Y09764; CAA70903.1; JOINED.
DR EMBL: Y09765; CAA70904.1; -
DR EMBL: Y07637; CAA68914.1; -
DR EMBL: U92283; AAB94645.1; -
DR EMBL: U92281; AAB94645.1; JOINED.
DR EMBL: U92282; AAB94645.1; JOINED.
DR GeneW: HGNC:4085; GABRE.
DR MIM: 300093; -
DR GO: GO:0005887; C: integral to plasma membrane; TAS.
DR GO: GO:0004890; F: GABA-A receptor activity; TAS.
DR GO: GO:0007214; P: gamma-aminobutyric acid signaling pathway; TAS.
DR GO: GO:0006832; P: small molecule transport; TAS.
DR InterPro: IPR006029; Neu_channel_memb.
DR InterPro: IPR006202; Neu_channel_LBD.
DR InterPro: IPR006201; Neu_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR TIGRfams: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
DR Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KM Transmembrane.
KW SIGNAL
FT SIGNAL 1 22
FT CHAIN 23 506
FT DOMAIN 23 253
FT TRANSMEM 254 274
FT TRANSMEM 281 301
FT TRANSMEM 344 364
FT DOMAIN 365 485
FT TRANSMEM 486 506
FT CARBOHYD 134 134
FT CARBOHYD 252 252
FT DISULFID 195 209
FT CONFLICT 102 102
FT CONFLICT 245 245
FT CONFLICT 261 261
FT CONFLICT 471 471
FT CONFLICT 502 502
SQ SEQUENCE 506 AA; 57971 MW; 013EE136ECCFC73 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 506;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNMC 6
DB 458 CMAPDC 463

RESULT 15

ADDEC_ARCFU STANDARD; PRT; 556 AA.
AC 029999;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable adenine deaminase (EC 3.5.4.2) (Adenase) (Adenine aminase).
GN Af0240.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftis B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: Adenine + H(2)O = hypoxanthine + NH(3).
CC -1- SIMILARITY: BELONGS TO THE ADENINE DEAMINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AE001089; AAB90993.1; -
DR FIR: H62279; H69279.
DR TIGR: AF0240; -
DR InterPro: IPR006679; Adenine_deam.
DR InterPro: IPR006680; Amidohydro_1.
DR InterPro: IPR005847; Urease.
DR Pfam: PF01979; Amidohydro_1; 1.
DR Pfam: PF000518; Urease; 1.
DR TIGRfams: TIGR01178; ade; 1.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 556 AA; 60764 MW; B3E3440A09F86A79 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 556;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNMC 6
DB 137 CMAPSC 142

Search completed: September 4, 2003, 21:01:47
Job time: 5.20588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 24.6176 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-63
Perfect score: 39
Sequence: 1 CVPNC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	97.4	129	10	08h7y2
2	38	97.4	135	10	040228
3	38	97.4	135	10	043538
4	36	92.3	303	16	09hvr7
5	35	89.7	288	5	09xyv5
6	35	89.7	334	5	024403
7	35	89.7	334	5	09vaw8
8	35	89.7	366	15	0997p7
9	35	89.7	472	10	09m1p3
10	35	89.7	505	16	09ksf2
11	35	89.7	507	16	09hmf8
12	35	89.7	507	16	08dae0
13	35	89.7	613	5	09vgr8
14	34	87.2	197	10	0944d8
15	34	87.2	308	11	08clt7
16	34	87.2	335	13	08jful

17	34	87.2	353	4	08ly68	08ly68 homo sapien
18	34	87.2	445	5	08t364	08t364 podocoryne
19	34	87.2	698	5	08la01	08la01 plasmidom
20	34	87.2	712	5	020323	020323 caenorhabdi
21	34	87.2	772	12	08jrs8	08jrs8 diachasmimo
22	34	87.2	1403	11	0922d3	0922d3 mus musculu
23	34	87.2	1898	11	08chr2	08chr2 mus musculu
24	34	87.2	2192	11	0922u2	0922u2 mus musculu
25	33	84.6	73	6	09n052	09n052 macaca fasc
26	33	84.6	86	9	094m96	094m96 lactococcus
27	33	84.6	99	4	09ujn6	09ujn6 homo sapien
28	33	84.6	105	4	060801	060801 homo sapien
29	33	84.6	105	4	096900	096900 homo sapien
30	33	84.6	106	13	090xy4	090xy4 brachydanio
31	33	84.6	126	6	09n0b8	09n0b8 macaca fasc
32	33	84.6	144	4	08nbt2	08nbt2 homo sapien
33	33	84.6	200	12	09ict0	09ict0 porcine rep
34	33	84.6	200	12	041198	041198 porcine rep
35	33	84.6	200	12	041196	041196 porcine rep
36	33	84.6	200	12	041193	041193 porcine rep
37	33	84.6	200	12	041192	041192 porcine rep
38	33	84.6	200	12	041195	041195 porcine rep
39	33	84.6	200	12	041195	041195 porcine rep
40	33	84.6	200	12	041188	041188 porcine rep
41	33	84.6	200	12	041197	041197 porcine rep
42	33	84.6	200	12	084936	084936 porcine rep
43	33	84.6	200	12	09j735	09j735 porcine rep
44	33	84.6	201	2	09anb9	09anb9 bradyrhizob
45	33	84.6	234	11	08k0t6	08k0t6 mus musculu

ALIGNMENTS

RESULT 1

08h7y2 PRELIMINARY: PRT; 129 AA.
ID 08h7y2
AC 08h7y2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OJ1607A12.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaristaceae; Oryzaceae; Oryza.
OX NCBL_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA Currie J., Collura K.,
RT "Rice Genomic Sequence."
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105729; AAN06866.1; -
KW Hypothetical protein.
SQ SEQUENCE 129 AA; 13261 MW; 775FE060768DAF82 CRC64;

Query Match 97.4%; Score 38; DB 10; Length 129;
Best Local Similarity 83.3%; Pred. No. 2.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPNC 6
DB 93 CVPNC 98

RESULT 2

ID 040228 PRELIMINARY: PRT; 135 AA.
AC 040228;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 GN ORF (Fragment).
 OS Liliaceae.
 OC Liliaceae.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Liliaceae; Liliaceae;
 RN NCB1_TaxID=4690;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Flower buds;
 RA Tabata S.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Flower buds;
 RA MEDLINE=96051386; PubMed=7584025;
 RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A.,
 RT "Characterization of cDNAs induced in meiotic prophase in Lily
 RT Microsporocytes.";
 RL DNA Res. 1:15-26(1994).
 DR EMBL; D21811; BAA04835.1;
 FT NON_TER
 SO SEQUENCE 135 AA; 14769 MW; D7E0AE91EFCDAE6B CRC64;
 Query Match
 Best Local Similarity 97.4%; Score 38; DB 10; Length 135;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CVPNC 6
 Db 130 CVIPLNC 135

DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 GN ORF (Fragment).
 OS Liliaceae.
 OC Liliaceae.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Liliaceae; Liliaceae;
 RN NCB1_TaxID=4690;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Flower buds;
 RA Tabata S.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Flower buds;
 RA MEDLINE=96051386; PubMed=7584025;
 RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A.,
 RT "Characterization of cDNAs induced in meiotic prophase in Lily
 RT Microsporocytes.";
 RL DNA Res. 1:15-26(1994).
 DR EMBL; D21811; BAA04835.1;
 FT NON_TER
 SO SEQUENCE 135 AA; 14769 MW; D7E0AE91EFCDAE6B CRC64;
 Query Match
 Best Local Similarity 97.4%; Score 38; DB 10; Length 135;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CVPNC 6
 Db 130 CVIPLNC 135

RESULT 4
 ID 09HVR7 PRELIMINARY; PRT; 303 AA.
 AC 09HVR7;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DE Probable permease of ABC transporter.
 GN PA4504.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 RN NCB1_TaxID=287;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RA MEDLINE=20437337; PubMed=10984043;
 RA Hickey M.J., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Smith K.A., Spencer D.H., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT Opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004864; AAG07892.1;
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 SO Complete proteome.
 SEQUENCE 303 AA; 32367 MW; B9496E9EA5E82B3 CRC64;
 Query Match
 Best Local Similarity 92.3%; Score 36; DB 16; Length 303;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CVPNC 6
 Db 211 CVIPLNC 216

RESULT 5
 ID 09XIV5 PRELIMINARY; PRT; 288 AA.
 AC 09XIV5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
 DE C-type lectin Tc-cl-4.
 OS Toxocara canis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Toxocaridae; Toxocara.
 RN NCB1_TaxID=6265;
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21011334; PubMed=11128806;
 RA Loukas A., Doedens A., Hintz M., Matzels R.M.;
 RT "Identification of a new C-type lectin, TcS-70, secreted by infective
 RT larvae of Toxocara canis, which binds to host ligands.";
 RL Parasitology 121:545-554(2000).
 DR EMBL; AF126830; AAD31000.1;
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF00059; Lectin_C; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 SO SEQUENCE 288 AA; 30273 MW; 5C08AA09A1311064 CRC64;
 Query Match
 Best Local Similarity 89.7%; Score 35; DB 5; Length 288;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
 11111
 Db 120 CAVPNC 125

RESULT 6

ID 024403 PRELIMINARY: PRT: 334 AA.
 AC 024403:
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE Sperm protein.
 GN M5798CA OR CG11719.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGONR;
 RA Schaefer M., Boersch D., Huelster A., Schaefer U.;
 RL Submitted (Aug-1992) to the EMBL/Genbank/DBJ databases.
 DR EMBL: X67704; CAA47941.1; -;
 DR HSSP: P01180; INPO.
 DR FlyBase: FBgn0002865; M5798CA.
 DR InterPro: IPR001450; 4Fe4s_ferredoxin.
 DR InterPro: IPR001368; TNFR_C6.
 DR PROSITE: PS00198; 4Fe4s_FERREDOXIN; 3.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 SQ SEQUENCE 334 AA; 34709 MW; 2C8361C34BB9CGF CRC64;

Query Match 89.7% Score 35; DB 5; Length 334;
 Best Local Similarity 83.3% Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
 11111
 Db 242 CAVPNC 247

RESULT 7

ID 09VAM8 PRELIMINARY: PRT: 334 AA.
 AC 09VAM8:
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE M5798CA protein (GH09231P).
 GN M5798CA OR CG11719.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Medline-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Caveley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajal M., Kalush E., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jammal B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimmos I., Simpson M., Skipski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster."
 RT Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Fatih D., Flise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE003764; AA56779.1; -;
 DR EMBL: AY060667; AAL28215.1; -;
 DR HSSP: P01180; INPO.
 DR FlyBase: FBgn0002865; M5798CA.
 DR InterPro: IPR001450; 4Fe4s_ferredoxin.
 DR InterPro: IPR001368; TNFR_C6.
 DR PROSITE: PS00198; 4Fe4s_FERREDOXIN; 3.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 SQ SEQUENCE 334 AA; 34725 MW; 2C876D1864BFC94 CRC64;

Query Match 89.7% Score 35; DB 5; Length 334;
 Best Local Similarity 83.3% Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6
 11111
 Db 242 CAVPNC 247

RESULT 8

ID 0997P7 PRELIMINARY: PRT: 366 AA.
 AC 0997P7:
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DE 01-OCT-2002 (TREMBLrel. 22, last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 CC Viruses; Retroviruses; Retroviridae; Lentivirus.
 CC NCB1_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang F., Kimura T., Nishihara K., Yoshimura K., Koito A.,
 RA Matsushita S.;
 RA "Emergence of autologous neutralization-resistant variants from pre-
 RT exisitant quasiparticles during viral-rebound from patients under
 RT HART."
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AB059310; BAB40894.1; -
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT 366
 SQ SEQUENCE 366 AA; 40154 MW; 2B7683B308FC8CB8 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 15; Length 366;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVPNC 6
 Db 1 CVPNC 6

RESULT 9
 ID 09MIP3 PRELIMINARY; PRT; 472 AA.

AC 09MIP3; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Protein kinase-like protein.
 GN t18B2.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordan N., Bangert S., Wiedelmann R., Voss H., Unseld M., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Queitner F., Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC EMBL: AL136652; CAB72463.1; -
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 2.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 472 AA; 52798 MW; 583042CFB904A3B3 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 10; Length 472;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPNC 6
 Db 268 CMIPNC 273

RESULT 10
 ID 09KSF2 PRELIMINARY; PRT; 505 AA.
 AC 09KSF2; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Fumarate hydratase, class I, putative.
 GN VC1304.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Rasmussen M.D., Vamathevan J., Bess S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith R., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004210; AAF94463.1; -
 DR TIGR: VC1304; -
 DR InterPro: IPR004646; TlGA_fumA_fumb.
 DR InterPro: IPR004647; TlGB_fumA_fumb.
 DR TIGRFAMs: TIGR00722; tlGA_fumA_fumb; 1.
 DR TIGRFAMs: TIGR00723; tlGB_fumA_fumb; 1.
 KW Complete proteome.
 SQ SEQUENCE 505 AA; 54778 MW; 40C4A2C06449D41F CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 16; Length 505;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPNC 6
 Db 272 CMIPNC 277

RESULT 11
 ID 09HW68 PRELIMINARY; PRT; 507 AA.
 AC 09HW68; 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Probable fumarate.
 GN PA433.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004849; AAC07721.1; -
 DR InterPro: IPR004646; TlGA_fumA_fumb.
 DR InterPro: IPR004647; TlGB_fumA_fumb.
 DR TIGRFAMs: TIGR00722; tlGA_fumA_fumb; 1.
 DR TIGRFAMs: TIGR00723; tlGB_fumA_fumb; 1.
 KW Complete proteome.
 SQ SEQUENCE 507 AA; 54763 MW; 49BCDFA9F172907 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 16; Length 507;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 272 CIMPNC 277

RESULT 12

OBDAEO PRELIMINARY: PRT: 507 AA.

AC 08DAEO: 0

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Fumarate hydratase.

GN VVI2266.

OS Vldio vulnificus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vldrio.

OX NCBI_TaxID=672;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CMCP6;

RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,

RA Choy H.E.;

RT "Complete genome sequence of *Vibrio vulnificus* CMCP6";

RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL: AE016804; AA010644.1; -.

KW Complete proteome.

SO SEQUENCE 507 AA; 54848 MW; 210DF532885F2B7 CRC64;

Query Match 89.7%; Score 35; DB 16; Length 507;

Best Local Similarity 66.7%; Pred. No. 36;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPNC 6

DB 272 CIMPNC 277

RESULT 13

Q9VGR8 PRELIMINARY: PRT: 613 AA.

AC Q9VGR8;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE CG6689 protein.

GN CG6689.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydriidae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Cealinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abail J.F., Agdayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck H., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durtin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,

RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostreli A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Relvert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*";

RL Science 287:2185-2195(2000).

DR EMBL: AE003690; AAF54607.1; -.

DR Flybase: FBgn0037877; CG6689.

DR InterPro: IPR006612; DUF.DM3.

DR InterPro: IPR007087; Znf_C2H2.

DR InterPro: IPR006025; Zn_MTPpeptase.

DR Pfam: PF00096; zf-C2H2; 7.

DR ProDom: PD000003; Znf_C2H2; 1.

DR SMART: SM00692; DM3; 1.

DR SMART: SM00355; Znf_C2H2; 7.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.

DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 7.

DR PROSITE: PS00142; ZINC_PROTEASE; 1.

KW Metal-binding; Zinc; Zinc-finger.

SO SEQUENCE 613 AA; 70431 MW; 8C9AD2899C108B4 CRC64;

Query Match 89.7%; Score 35; DB 5; Length 613;

Best Local Similarity 83.3%; Pred. No. 43;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6

DB 3 CAVPNC 8

RESULT 14

Q94AD8 PRELIMINARY: PRT: 197 AA.

AC Q94AD8;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Resistance disease protein.

GN TUR7.

OS Brassica oleracea (Cauliflower).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI_TaxID=3712;

RN [1]

RP SEQUENCE FROM N.A.

RA Cao B., Lei J., Song H.;

RT "Tur7-resistance disease gene in *Brassica oleracea*";

RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF432595; AML28107.1; -.

SO SEQUENCE 197 AA; 22932 MW; AE90A8E2C23F854 CRC64;

Query Match 87.2%; Score 34; DB 10; Length 197;

Best Local Similarity 83.3%; Pred. No. 25;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPNC 6

DB 26 CVPNC 31

RESULT 15

08C1T7 PRELIMINARY; PRT; 308 AA.
 AC 08C1T7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=spinal cord;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT *Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.*;
 RL Nature 420:563-573(2002).
 DR EMBL; AK090368; BAC41186.1; -.
 KW Hypothetical protein.
 FT NON_TER 308
 FT NON_TER 308
 SO SEQUENCE 308 AA; 34483 MW; AD58F34417BB513D CRC64;

Query Match

87.2%; Score 34; DB 11; Length 308;

Best Local Similarity 83.3%; Pred. No. 37;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPNC 6

11111
 Db 84 CVPNC 89

Search completed: September 4, 2003, 21:06:48
 Job time : 25.7086 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 36.0294 Seconds
(Without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893D-64
Perfect score: 43
Sequence: 1 CAVVNC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	23	AAE16726
2	43	100.0	7	23	AAE16729
3	37	86.0	7	23	AAE16715
4	37	86.0	7	23	AAE16718
5	36	83.7	124	22	AAU48588
6	35	81.4	155	23	ABP42962
7	34	79.1	8	23	AAE16727
8	34	79.1	8	23	AAE16730
9	34	79.1	67	21	AAE25741

10	34	79.1	78	22	AAW91272	Human immune/haema
11	34	79.1	79	22	AAW91271	Human immune/haema
12	34	79.1	119	22	ABG09232	Novel human diaeno
13	34	79.1	156	23	ABBB9784	Human polypeptide
14	34	79.1	235	23	AAE25742	Human secreted pro
15	34	79.1	316	23	AAE22089	Human novel ion ex
16	34	79.1	331	21	AAE1847	Human ORFX ORF1611
17	34	79.1	344	21	AAE19140	Human secreted pro
18	34	79.1	353	23	AAE22090	Human novel ion ex
19	34	79.1	420	23	AAE14454	Human protein phos
20	34	79.1	428	22	AAE80228	Human PRO269 prote
21	34	79.1	431	22	AAE85042	Shrimp white spot
22	34	79.1	446	23	ABP43556	Human PRO269 prote
23	34	79.1	480	23	ABG12623	Novel human diaeno
24	34	79.1	490	20	AAV13360	Amino acid sequenc
25	34	79.1	490	21	AAE33460	Human PRO269 prote
26	34	79.1	490	21	AAE24403	Human PRO269 prote
27	34	79.1	490	21	AAV95016	Human secreted pro
28	34	79.1	490	22	AAU12349	Human PRO269 polyp
29	34	79.1	490	22	AAE06596	Human protein havi
30	34	79.1	490	23	AAU66134	Human PRO269 polyp
31	34	79.1	490	24	ABR47831	Human secreted pro
32	34	79.1	490	24	ABU69638	Novel human secret
33	34	79.1	490	24	ABU71461	Human PRO polypept
34	34	79.1	490	24	ABU71907	Human secreted/tra
35	34	79.1	490	24	ABU66747	Human PRO polypept
36	34	79.1	490	24	ABU67023	Human secreted/tra
37	34	79.1	490	24	ABU67361	Human secreted pro
38	34	79.1	490	24	ABU59828	Novel secreted and
39	34	79.1	490	24	ABU64515	Human secreted/tra
40	34	79.1	490	24	ABU54363	Human novel ion ex
41	34	79.1	603	23	AAE22088	Human secreted/tra
42	34	79.1	647	24	ABU56572	Lung cancer-associa
43	34	76.7	42	13	AAE25686	Exon IX of human h
44	34	76.7	86	21	AAE24597	Human secreted pro
45	33	76.7	87	21	AAE24487	Human secreted pro

ALIGNMENTS

RESULT 1	AAE16726	standard; peptide: 7 AA.
XX	AAE16726:	
AC	AAE16726:	
DT	09-APR-2002 (first entry)	
XX		
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPV derive peptide #6.	
XX		
KW	Hydrolysed elastin peptide: HEP, therapy; cosmetic appearance; cardiant;	
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;	
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;	
KW	hypotensive; chronic obstructive pulmonary disease; antihypertensive;	
KW	restenosis post-angioplasty; antidiabetic; antidiabetic; thrombolytic.	
XX		
OS	Ligamentum nuchae.	
OS	Synthetic.	
PN	W0200191700-A2.	
XX		
PD	06-DEC-2001.	
XX		
PF	30-MAY-2001; 2001WO-US17384.	
XX		
PR	30-MAY-2000; 2000US-0580110.	
PR	30-MAY-2000; 2000US-0580156.	
PR	30-MAY-2000; 2000US-0580893.	
PR	30-MAY-2000; 2000US-0584001.	
XX		
XX		
PA	(CONN-) CONNECTIVE TISSUE IMAGINGING LLC.	
XX		

PI	Mitts TF, Sandberg LB, Jimenez F;
XX	
DR	WPI: 2002-106259/14.
XX	
PT	Composition for improving the elasticity of tissue comprises peptide
XX	
PS	substances which mimic the sequences found in elastin -
XX	
XX	Claim 24; Page 21; 53pp; English.
CC	
CC	The present invention relates to a composition for the treatment of
CC	mammalian tissue comprising peptides (hydrolysed elastin peptide, HEP) or
CC	their biological equivalents. The composition is useful not only for
CC	improving the cosmetic appearance of skin, but for the treatment of
CC	various skin disorders associated with reduced elastin formation and for
CC	numerous other tissue disorders especially of blood vessels including
CC	hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC	thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC	angioplasty. They may also be useful in coating surgical appliances i.e.
CC	stents. The use of peptides and peptide fragments which mimic those found
CC	in elastin overcomes the problem of solubility of the dense cross-linked
CC	structure of elastin itself, which renders it difficult to use
CC	therapeutically. The peptides are found to improve elastin production in
CC	treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC	elastin peptide (HEP), VVPN derived peptide.
XX	
SQ	Sequence 7 AA:
	Query Match 100.0%; Score 43; DB 23; Length 7;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CAVVNC 7 1 CAVVNC 7
Db	
	RESULT 2
	AAL616729
ID	AAL616729 standard; peptide: 7 AA.
XX	
AC	AAL616729;
DT	09-APR-2002 (first entry)
XX	
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #9.
XX	
KW	Hydrolysed elastin peptide; HBP; therapy; cosmetic appearance; cardiant;
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypertensive; chronic obstructive pulmonary disease; antlerteriosclerotic;
KW	restenosis post-angioplasty; antiangular; anticoagulant; thrombolytic;
KW	cyclic.
OS	Ligamentum nuchae.
XX	
XX	Synthetic.
FT	Key
FT	Modified-site 1 Location/Qualifiers
FT	/note= "This residue forms a cyclic structure with Cys
FT	at position 7 with copper as a chelating agent"
FT	Modified-site 7
FT	/note= "This residue forms a cyclic structure with Cys
FT	at position 1 with copper as a chelating agent"
PN	WO200191700-A2.
XX	
PD	06-DEC-2001.
XX	
PX	30-MAY-2001; 2001WO-US17384.
PR	30-MAY-2000; 2000US-0580110.
PR	30-MAY-2000; 2000US-0580156.
PR	30-MAY-2000; 2000US-0580893.
RR	

PR	30-MAY-2000; 2000US-0584001.
XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGING/ENGINEERING LLC.
XX	
PI	Milts TF, Sandberg LB, Jimenez F;
XX	
DR	WPI: 2002-106259/14.
XX	
PT	Composition for improving the elasticity of tissue comprises peptide
PS	substances which mimic the sequences found in elastin
XX	
XX	Claim 24; Page 21; 53pp: English.
CC	
CC	The present invention relates to a composition for the treatment of
CC	mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
CC	their biological equivalents. The composition is useful not only for
CC	improving the cosmetic appearance of skin, but for the treatment of
CC	various skin disorders associated with reduced elastin formation and for
CC	numerous other tissue disorders especially of blood vessels including
CC	hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC	thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC	angioplasty. They may also be useful in coating surgical appliances i.e.
CC	stents. The use of peptides and peptide fragments which mimic those found
CC	in elastin overcomes the problem of solubility of the dense cross-linked
CC	structure of elastin itself, which renders it difficult to use
CC	therapeutically. The peptides are found to improve elastin production in
CC	treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC	elastin peptide (HEP), VVPM derived cyclic peptide.
XX	
SO	Sequence 7 AA:
	Query Match 100.0%; Score 43; DB 23; Length 7;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CAVVNC 7
DB	1 CAVVNC 7
	RESULT 3
	AAE16715
ID	AAE16715 standard; peptide; 7 AA.
XX	
AC	AAE16715;
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPM derived peptide #9.
XX	
KW	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW	restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
KW	cyclic.
XX	
OS	Ligamentum nuchae.
OS	Synthetic.
XX	
FH	Key
FT	Disulfide-bond 1..7
XX	
PN	Location/Qualifiers
XX	
PD	WO200191700-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-US17384.
XX	
PR	30-MAY-2000; 2000US-0580110.
PR	30-MAY-2000; 2000US-0580156.
PR	30-MAY-2000; 2000US-0580893.
PR	30-MAY-2000; 2000US-0584001.
PR	

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI WPI: 2002-106259/14.
 DR
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 20; 53pp: English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.
 CC
 XX Sequence 7 AA:
 SQ
 Query Match 86.0%; Score 37; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CAVVPC 7
 Db 1 CAVVPC 7
 RESULT 4
 ID AAE16718 standard; peptide: 7 AA.
 AC AAE16718;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPQ derived peptide #12.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy: cosmetic appearance; cardiant;
 KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypertensive; chronic obstructive pulmonary disease; arteriosclerosis;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 XX
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 1
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 7 with copper as a chelating agent"
 FT Modified-site 7
 FT /note="This residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 FT
 XX WO200191700-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 30-MAY-2001; 2001WO-US17384.
 XX

PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI WPI: 2002-106259/14.
 DR
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin
 XX
 PS Claim 24; Page 20; 53pp: English.
 XX
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VVPQ derived cyclic peptide.
 CC
 XX Sequence 7 AA:
 SQ
 Query Match 86.0%; Score 37; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CAVVPC 7
 Db 1 CAVVPC 7
 RESULT 5
 ID AAU48588 standard; Protein: 124 AA.
 AC AAU48588;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9484.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 OS
 XX
 FT WO200181581-A2.
 FT
 FT 01-NOV-2001.
 FT
 FT 20-APR-2001; 2001WO-US12865.
 FT
 FT 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 9783; 1069pp; English.
 XX
 CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), warts and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 124 AA:
 Query Match 83.7%; Score 36; DB 22; Length 124;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
 111111
 DB 87 CAVIPRC 93

RESULT 6
 ABA42962
 ID ABA42962 standard; Protein; 155 AA.
 XX
 AC ABA42962;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HPDVO67, SEQ ID NO:4094.
 XX
 KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 KW
 XX Homo sapiens.
 OS
 PN WO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.

XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Blise CE, Rosen CA;
 XX
 DR WPI: 2002-147878/19.
 DR N-PSDB; ABQ56039.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID No 4094; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABA41054-
 CC ABA43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 908 identical and polynucleotides 928 identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 155 AA:
 Query Match 81.4%; Score 35; DB 23; Length 155;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAVPNC 7
 111111
 DB 101 CAVIPHC 107

RESULT 7
 AAE16727
 ID AAE16727 standard; peptide; 8 AA.
 XX
 AC AAE16727;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic.
 XX
 OS Ligamentum nuchae.

XX	Synthetic.
XX	WO200191700-A2.
PN	
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-US17384.
XX	
PR	30-MAY-2000; 2000US-0580110.
XX	
PR	30-MAY-2000; 2000US-0580156.
XX	
PR	30-MAY-2000; 2000US-0580893.
XX	
PA	30-MAY-2000; 2000US-0584001.
XX	
(CONN-)	CONNECTIVE TISSUE IMAGING LLC.
XX	
P1	Mitts TF, Sandberg LB, Jimenez F;
XX	
DR	WPI; 2002-106259/14.
XX	
P7	Composition for improving the elasticity of tissue comprises peptide
XX	
P7	substances which mimic the sequences found in elastin
XX	
PS	Claim 24; Page 21; 53pp; English.
XX	
CC	The present invention relates to a composition for the treatment of
CC	mammalian tissue comprising peptides (hydrolysed elastin peptide, HEP) or
CC	their biological equivalents. The composition is useful not only for
CC	improving the cosmetic appearance of skin, but for the treatment of
CC	various skin disorders associated with reduced elastin formation and for
CC	hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC	thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC	angioplasty. They may also be useful in coating surgical appliances i.e.
CC	stents. The use of peptides and peptide fragments which mimic those found
CC	in elastin overcomes the problem of solubility of the dense cross-linked
CC	structure of elastin itself, which renders it difficult to use
CC	therapeutically. The peptides are found to improve elastin production in
CC	treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC	elastin peptide (HEP), VVPN derived peptide.
XX	
SQ	Sequence 8 AA;
OY	Query Match 79.1%; Score 34; DB 23; Length 8;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dd	2 AVVPMC 7 3 AVVPMC 8
RESULT 8	
AHEL6730	
ID	AHEL6730 standard; peptide; 8 AA.
XX	
AC	AHEL6730;
XX	
DT	09-APR-2002 (first entry)
XX	
Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.	
XX	
KM	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KM	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KM	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KM	hypertensive; chronic obstructive pulmonary disease; antierosclerotic;
KM	restenosis post-angioplasty; antianginal; anticoagulant; thrombolytic;
KM	cyclic.
OS	
Ligamentum nuchae.	
OS	
Synthetic.	
XX	
Key	Location/Qualifiers
FT	Modified-site 1

```

FT      /note= "This residue forms a cyclic structure with Cys
FT      at position 8 with copper as a chelating agent"
FT      Modified-site
FT      8
FT      /note= "This residue forms a cyclic structure with Cys
FT      at position 1 with copper as a chelating agent"
PM      WO200191700-A2.
XX
XX      PD      06-DEC-2001.
XX      PE      30-MAY-2001; 2001WO-US17384.
XX      PF      30-MAY-2000; 2000US-0580110.
XX      PG      30-MAY-2000; 2000US-0580136.
XX      PH      30-MAY-2000; 2000US-0580893.
XX      PI      30-MAY-2000; 2000US-0584001.
XX
XX      PA      (CONN-) CONNECTIVE TISSUE IMAGINERLING LLC.
XX      PT      Miltis TF, Sandberg LB, Jimenez F;
XX      PU      WPI; 2002-106259/14.
XX      PR      Composition for improving the elasticity of tissue comprises peptide
XX      PT      substances which mimic the sequences found in elastin
XX      PS      Claim 24; Page 21; 53pp; English.
XX
XX      The present invention relates to a composition for the treatment of
XX      mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
XX      their biological equivalents. The composition is useful not only for
XX      improving the cosmetic appearance of skin, but for the treatment of
XX      various skin disorders associated with reduced elastin formation and for
XX      numerous other tissue disorders especially of blood vessels including
XX      hypertension, coronary heart disease, arteriosclerosis, angina, coronary
XX      thrombosis, chronic obstructive pulmonary disease and restenosis post-
XX      angioplasty. They may also be useful in coating surgical appliances i.e.
XX      stents. The use of peptides and peptide fragments which mimic those
XX      in elastin overcomes the problem of solubility of the dense cross-linked
XX      structure of elastin itself, which renders it difficult to use
XX      therapeutically. The peptides are found to improve elastin production in
XX      treated tissues. The present sequence is ligamentum nuchae hydrolysed
XX      elastin peptide (HEP), VVPN derived cyclic peptide.
XX
XX      SO      Sequence 8 AA:
XX
XX      Query Match          79.1%; Score 34; DB 23; Length 8;
XX      Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
XX      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      2 AVVENC 7
XX      |||||
XX      DB      3 AVVENC 8
XX
XX      RESULT 9
XX      AAB25741
XX      ID      AAB25741 standard; Protein: 67 AA.
XX
XX      AAB25741;
XX
XX      04-DEC-2000 (first entry)
XX
XX      DE      Human secreted protein sequence encoded by gene 16 SEQ ID NO:130.
XX
XX      Human; secreted protein; immunosuppressive; immunostimulant; nootropic;
XX      antiinflammatory; cardiact; vulnerary; antitumor; anticonvulsant;
XX      antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
XX      antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
XX      gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
XX      cancer; immune system disorder; hyperproliferative disorder; infection;
XX      cardiovascular disorder; neurological disease; wound healing.
XX

```

OS Homo sapiens.
 XX WO200043495-A2.
 XX 27-JUL-2000.
 PD
 XX 18-JAN-2000; 2000WO-US00903.
 XX 19-JAN-1999; 99US-0116330.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Edner R, Young PE, Ni J, Moore PA;
 PI Komatsoulis G, Birse CE;
 XX
 DR WPI; 2000-499225/44.
 XX
 PT New isolated polynucleotide encoding a secreted protein useful for
 PT preventing, treating or ameliorating a medical condition -
 XX
 PS Disclosure; Page 48; 451p; English.
 XX
 CC The polynucleotide sequences given in AAB87666 to AAB87708 encodes the
 CC human secreted proteins given in AAB25665 to AAB25755. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC immunostimulant; antiinflammatory; cardiant; vulnary; antilicer;
 CC neurotropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
 CC antibacterial; antiparasitic; thrombolytic; anticonstant;
 CC antihypertensive and cytostatic. The secreted proteins and their
 CC polynucleotides can be used in gene therapy and as vaccines,
 CC chemotaxis-modulators and angiogenesis- modulators. The human secreted
 CC proteins and polynucleotides can be used for diagnosing (the
 CC susceptibility to) a pathological condition by determining the presence
 CC of absence or amount of expression of the protein. The polynucleotides and
 CC proteins can also be used in the treatment and diagnosis of cancer,
 CC diseases of the immune system, hyperproliferative disorders,
 CC cardiovascular disorders and neurological disease. They can also be used
 CC to promote wound healing and to fight infection. AAB87657 to AAB87665 and
 CC AAB85664 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 67 AA;
 Query Match 79.1%; Score 34; DB 21; Length 67;
 BestLocal Similarity 57.1%; Pred. No. 51;
 Matches 4; Conservative. 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CAVVENC 7
 1 :111
 52 CVTIPNC 58
 DB
 RESULT 10
 ID AAM91272 standard; Protein; 78 AA.
 AC AAM91272;
 XX
 XX 07-NOV-2001 (first entry)
 DT
 XX
 DE Human Immune/haematopoietic antigen SEQ ID NO:18865.
 XX
 XX Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;
 KM cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 XX WO200157182-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX

PF 17-JAN-2001; 2001WO-US01354.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234423.
 PR 21-SEP-2000; 2000US-0234474.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.

PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236602
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	01-NOV-2000	2000US-0241826
PR	08-NOV-2000	2000US-0246167
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249264
PR	17-NOV-2000	2000US-0249265
PR	17-NOV-2000	2000US-0249287
PR	17-NOV-2000	2000US-0249289
PR	17-NOV-2000	2000US-0249300
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0251030
PR	05-DEC-2000	2000US-0251088
PR	05-DEC-2000	2000US-0256719
PR	06-DEC-2000	2000US-0251479
PR	06-DEC-2000	2000US-0251486
PR	08-DEC-2000	2000US-0251689
PR	08-DEC-2000	2000US-0251689
PR	08-DEC-2000	2000US-0251989
PR	08-DEC-2000	2000US-0251989
PR	11-DEC-2000	2000US-0254097
PR	03-JAN-2001	2001US-0259678
(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA, Barash SC, Ruben SM,	
XX	WPI, 2001-483426/52.	
XX	N-PSDB; AAK64053.	

PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT	useful for preventing, diagnosing and/or treating cancers and
PT	metastasis -
PS	xx
PS	Claim 11: SEQ ID NO 18865; 3071pp + Sequence Listing; English.
CC	xx
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC	activity, and can be used in gene therapy and vaccine production. (I)
CC	actively, and polynucleotides may be used in the prevention, diagnosis and
CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	supplement the patients own production of (I). Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I), by inserting the
CC	nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/hematopoietic-related diseases, especially
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC	to AAK87694 represent human immune/hematopoietic antigen genomic
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC	represent sequences used in the exemplification of the present invention.
XX	xx
SQ	Sequence 78 AA:
Query Match	79.1% Score 34; DB 22; Length 78;
Best Local Similarity	57.1%; Pred. No. 59;
Matches 4; Conservative	2; Mismatches 1; Indels 0; Gaps 0.
OY	1 CAVVPCNC 7 :: 18 CLILPNC 24
Dd	
RESULT 11	
AAM91271	
ID	AAW91271 standard; Protein; 79 AA.
XX	
AC	AAW91271;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:18864.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis.
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
PD	09-AUG-2001.
XX	
PE	17-JAN-2001; 2001WO-US01354.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198122.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
N-PDB; AAK64052.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis

Claim 11; SEQ ID NO 18864; 3071bp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
suppress the patients' own production of (I). Additionally, (I)
polynucleotides may be used to produce the secreted (I), by inserting the
nucleic acids into a host cell and culturing the cell to express the
protein. (I) proteins and polynucleotides may be used to prevent,
diagnose and treat immune/hematopoietic-related diseases, especially

CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX

SO Sequence 79 AA;

Query Match 79.1%; Score 34; DB 22; Length 79;
 Best Local Similarity 57.1%; Pred. No. 60;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPC 7
 | : : : :
 Db 18 CLILPNC 24

RESULT 12

ABG09232
 ID ABG09232 standard; Protein: 119 AA.

AC ABG09232;

DF 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9223.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HXSE-) HXSE0 INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB: AAS73419.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 39591; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes.
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 119 AA;

Query Match 79.1%; Score 34; DB 22; Length 119;
 Best Local Similarity 57.1%; Pred. No. 88;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPC 7
 | : : : :
 Db 108 CKIMPNC 114

RESULT 13

ABB89784
 ID ABB89784 standard; Protein: 156 AA.

AC ABB89784;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 2160.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

PN WO200190304-A2.

PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI: 2002-122018/16.

DR N-PSDB: ABL90193.

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -

PS Claim 11; SEQ ID NO 2160; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 156 AA:

Query Match 79.1%; Score 34; DB 23; Length 156;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
111111
Db 123 CAVVNC 129

RESULT 14

AAB25742

ID AAB25742 standard; Protein; 235 AA.

AC AAB25742;

DT 04-DEC-2000 (first entry)

DE Human secreted protein sequence encoded by gene 16 SEQ ID NO:131.

Human: secreted protein; immunosuppressive; immunostimulant; neutrotropic;
antiinflammatory; cardiant; vulnery; antilcer; anticonvulsant; neutrotropic;
antiparkinsonian; neuroprotective; antiviral; antibacterial; cycostatic;
antiparasitic; thrombolytic; anticoagulant; antiatherosclerotic;
gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
cancer; immune system disorder; hyperproliferative disorder; infection;
cardiovascular disorder; neurological disease; wound healing.

OS Homo sapiens.

PN W0200043495-A2.

PD 27-JUL-2000.

PF 18-JAN-2000; 2000WO-US00903.

PR 19-JAN-1999; 99US-0116330.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Edner R, Young PE, Ni J, Moore PA;

PI Komatsoulis G, Birse CE;

DR WPI: 2000-499225/44.

PT New isolated polynucleotide encoding a secreted protein useful for

PT preventing, treating or ameliorating a medical condition -

PS Disclosure; Page 48; 451pp; English.

The polynucleotide sequences given in AAB25742 to AAB25742 encodes the
human secreted proteins given in AAB25665 to AAB25755. Human secreted
proteins have activities based on the tissues and cells the genes are
expressed in. Examples of activities include: immunosuppressive;
immunostimulant; antiinflammatory; cardiant; vulnery; antilcer;
neutrotropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
antibacterial; antiparasitic; thrombolytic; anticoagulant;
antiatherosclerotic and cycostatic. The secreted proteins and their
polynucleotides can be used in gene therapy and as vaccines,
chemotaxis-modulators and angiogenesis-modulators. The human secreted
proteins and polynucleotides can be used for diagnosing (the
susceptibility to) a pathological condition by determining the presence
of absence of a mutation in the polynucleotide or determining the
presence or amount of expression of the protein. The polynucleotides and
proteins can also be used in the treatment and diagnosis of cancer,
diseases of the immune system, hyperproliferative disorders,
cardiovascular disorders and neurological disease. They can also be used
to promote wound healing and to fight infection. AAB25657 to AAB25665 and
AAB25664 represent sequences used in the exemplification of the present
invention.

SQ Sequence 235 AA;

Query Match 79.1%; Score 34; DB 21; Length 235;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
111111
Db 52 CAVVNC 58

RESULT 15

AAE22089

ID AAE22089 standard; Protein; 316 AA.

AC AAE22089;

DT 25-JUL-2002 (first entry)

DE Human novel ion exchanger protein (NHP) #2.

Human: novel ion exchanger protein; NHP; gene mapping; forensic biology;
restriction fragment length polymorphism; RFLP analysis; chromosome 14.

OS Homo sapiens.

PN W0200226980-A2.

PD 04-APR-2002.

PF 24-SEP-2001; 2001WO-US29828.

PR 27-SEP-2000; 2000US-235745P.

PA (LEXI-) LEXICON GENETICS INC.

PI Friddle CJ, Gerhardt B;

DR WPI: 2002-362496/39.

DR N-PSDB; AAD35145.

PT New isolated nucleic acid molecule for encoding a novel human ion

PT exchanger protein -

PS Claim 1; Page 37-38; 40pp; English.

The invention relates to novel human ion exchanger proteins (NHP) and
their corresponding nucleic acid sequences. The nucleic acid sequence of
the invention is used for encoding a novel human ion exchanger protein,
and is useful for the identification of a protein coding sequence and
CC mapping a unique gene to a particular chromosome. It may also be used
CC as additional DNA markers for restriction fragment length polymorphism
CC (RFLP) analysis, and in forensic biology. The present sequence is human
NHP protein. NHP gene is located on chromosome 14.

SQ Sequence 316 AA;

Query Match 79.1%; Score 34; DB 23; Length 316;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
111111
Db 132 CAVVNC 138

Search completed: September 4, 2003, 21:00:36
Job time : 37.0294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 10.9118 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893D-64

Perfect score: 43

Sequence: 1 CAVVPMC 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	86.0	409	2 F90231	UDP-glucose 6-dehy
2	34	79.1	368	2 B70853	probable adhD prot
3	34	79.1	375	2 AB3618	alcohol dehydrogen
4	34	79.1	465	2 A56679	probable RNA-bind
5	34	79.1	644	2 T33132	hypothetical prote
6	34	79.1	784	2 J00317	hypothetical 82k p
7	33	76.7	127	2 S47895	SOH1 protein - yea
8	33	76.7	400	2 C70837	hypothetical prote
9	33	76.7	449	2 B71917	udp-n-acetylmuram
10	33	76.7	594	2 T23841	hypothetical prote
11	33	76.7	610	2 T18441	asparagine synthas
12	33	76.7	610	2 T23836	hypothetical prote
13	33	76.7	728	1 JH0579	hepatocyte growth
14	33	76.7	909	2 T06246	aspartate kinase (
15	33	76.7	916	2 T06242	aspartate kinase (
16	32	74.4	157	2 S41891	tryptophan 2-monoo
17	32	74.4	174	2 T11844	NAOH2 dehydrogenas
18	32	74.4	180	2 A45810	glycoprotein antiy
19	32	74.4	196	2 S73870	Mg308 homolog H10
20	32	74.4	379	2 S42543	hypothetical prote
21	32	74.4	519	2 D71881	phosphotransacyl
22	32	74.4	690	2 F70628	probable pta prote
23	32	74.4	755	1 DAGWT	tryptophan 2-monoo
24	32	74.4	755	1 OQAG4T	tryptophan 2-monoo
25	32	74.4	755	1 A13228	tryptophan 2-monoo
26	32	74.4	837	2 A34898	granulocyte colony
27	32	74.4	865	2 T41909	hypothetical prote
28	32	74.4	1014	2 T20361	hypothetical prote
29	31.5	73.3	1167	2 T06146	disease resistance

30	31	72.1	113	2 T24164	hypothetical prote
31	31	72.1	158	2 T26692	hypothetical prote
32	31	72.1	166	2 T28809	hypothetical prote
33	31	72.1	174	2 T14027	NAOH2 dehydrogenas
34	31	72.1	174	2 C59134	hypothetical prote
35	31	72.1	246	1 S29279	acetoacetyl-CoA re
36	31	72.1	274	2 D72044	hypothetical prote
37	31	72.1	274	2 A86581	hypothetical prote
38	31	72.1	334	2 AG2172	hypothetical prote
39	31	72.1	366	2 T39804	hypothetical prote
40	31	72.1	436	2 T28066	hypothetical prote
41	31	72.1	449	2 S76839	hypothetical prote
42	31	72.1	461	2 AB2185	hypothetical prote
43	31	72.1	469	2 T46929	hypothetical prote
44	31	72.1	469	2 T46930	hypothetical prote
45	31	72.1	511	2 JCT692	oligodendrocyte tr

ALIGNMENTS

RESULT 1
F90231
UDP-glucose 6-dehydrogenase (ugd) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: F90231
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aveyez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thl-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90231
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <KUR>
A:Cross-references: GB:AE006641; NID:g13813993; PIDN:AKK41109.1; GSPDB:GN00155
C:Genetics:
A:Gene: ugd
C:Superfamily: GDPmannose dehydrogenase

Query Match
Best Local Similarity 86.0%; Score 37; DB 2; Length 409;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
1 1 1 1 1 1
Db 213 CEVIPNC 219

RESULT 2
B70853
Probable adhD protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70853
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70853
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-368 <CO>
A:Cross-references: GB:AL021309; GB:ALJ23456; NID:g3261510; PIDN:CAI6144.1; PID:g280
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: adhD
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

F:25-359/Domain: long-chain alcohol dehydrogenase homology <LADH>

Query Match 79.1%; Score 34; DB 2; Length 368;
Best Local Similarity 57.1%; Pred. No. 54;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
I : : : : :
85 CSFVPMC 91

RESULT 3

AB3618
alcohol dehydrogenase (EC 1.1.1.1) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AB3618

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mufer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL54109.1; PID:g17985069; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:

A:Gene: BME110867
A:Map position: II
C:Superfamily: alcohol dehydrogenase: long-chain alcohol dehydrogenase homology

C:Keywords: oxidoreductase

Query Match 79.1%; Score 34; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
I : : : : :
93 CFFVPMC 99

RESULT 4

A56679
probable RNA-binding protein - house mosquito

C:Species: Culex pipiens (house mosquito)
C>Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jul-2000
C:Accession: A56679

R:Agarwal, M.; Bensadi, N.; Salgado, J.C.; Campbell, K.; Mouches, C.
Insect Biochem. Mol. Biol. 23, 621-629, 1993
A:Title: Characterization and genetic organization of full-length copies of a LINE retrovirus
A:Reference number: A56679; MUID:93357789; PMID:8394766
A:Accession: A56679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <AGAG>

A:Cross-references: GB:M91082; NTD:9403079; PIDN:AAA28290.1; PID:9403080

Query Match 79.1%; Score 34; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
I : : : : :
267 CAMVPMC 273

RESULT 5

T33132
hypothetical protein F41A4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33132
R:Bradshaw, H.; McPherson, C.

Submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid F41A4.

A:Reference number: Z21287

A:Accession: T33132

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-644 <BRAS>

A:Cross-references: EMBL:AF067610; PIDN:AACT17538.1; GSPDB:GN00022; CESP:F41A4.1

A:Experimental source: strain Bristol NZ; Clone F41A4

C:Genetics:

A:Gene: CESP:F41A4.1

A:Map position: 4

A:Introns: 57/3; 106/1; 147/2; 195/1; 244/2; 268/1; 308/2; 345/3; 374/3; 458/3

Query Match 79.1%; Score 34; DB 2; Length 644;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
I : : : : :
55 CSVPMC 61

RESULT 6

J00317
hypothetical 82K protein - Xanthomonas campestris pv. vesicatoria

C:Species: Xanthomonas campestris pv. vesicatoria

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993

C:Accession: J00317

R:Bonas, U.; Stail, R.E.; Staskawicz, B.

Mol. Gen. Genet. 218, 127-136, 1989

A:Title: Genetic and structural characterization of the avirulence gene avrBs3 from X

A:Reference number: J00316; MUID:89384426; PMID:2550761

A:Accession: J00317

A:Molecule type: DNA

A:Residues: 1-784 <BON>

A:Comment: X. campestris pv. vesicatoria is the causal agent of bacterial leaf spot d

C:Genetics:

A:Start codon: GTG

Query Match 79.1%; Score 34; DB 2; Length 784;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
I : : : : :
13 CLIPPMC 19

RESULT 7

S47895

SOH1 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein G2864; protein YGL127C

C:Species: Saccharomyces cerevisiae

C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 21-Jul-2000

C:Accession: S47895; S64138; S63786

R:Fan, H.; Klein, H.L.

Genetics 137, 945-956, 1994

A:Title: Characterization of mutations that suppress the temperature-sensitive growth

A:Reference number: S47895; MUID:95073608; PMID:7982575

A:Accession: S47895

A:Molecule type: DNA

A:Residues: 1-127 <FAN>

A:Cross-references: EMBL:L31921; NTD:9531011; PIDN:AAA35066.1; PID:9531012

R:Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64134

A:Accession: S64138

A:Molecule type: DNA

A:Residues: 1-127 <CER>

A:Cross-references: EMBL:Z72649; NTD:91322690; PID:g1322691; MIPS:YGL127C

A:Experimental source: strain 528C
R:Rodriguez-Belmonte, E.; Rodriguez-Torres, A.M.; Tizon, B.; Cadahia, J.L.; Gonzalez-Sis
Yeast 12, 145-148, 1996
A:Title: Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII reveals a
A:Reference number: S63783; MUID:96287651; PMID:866378
A:Accession: S63786
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-127 <ROD>
A:Cross-references: EMBL:X87252; NID:g1246906; PIDN:CA60704.1; PID:g1246910
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C:Genetics:
A:Gene: SCD:SOH1
A:Cross-references: SGD:S0003095; MIPS:YGL127C
A:Map position: 7L
C:Superfamily: Saccharomyces SOH1 protein
C:Keywords: transmembrane protein
F:78-94/Domain: transmembrane #status predicted <TM>

Query Match 76.7%; Score 33; DB 2; Length 127;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
1 1 1 1 1
Db 72 CAVVNC 78

RESULT 8
C70837
hypothetical protein RV0293c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70837
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajadurai, M.A.; Rogers, R.; Rutter, J.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:99295987; PMID:9634230
A:Accession: C70837
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-400 <CO>
A:Cross-references: GB:AL021930; GB:AL123456; NID:g3261524; PIDN:CA17368.1; PID:e125247
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0293c

Query Match 76.7%; Score 33; DB 2; Length 400;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AVVNC 7
1 1 1 1 1
Db 283 AVVNC 288

RESULT 9
B71917
udp-n-acetylmuramate-alanine ligase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: B71917
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: B71917
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-449 <ARN>
A:Cross-references: GB:AE001489; GB:AE001439; NID:g4155102; PIDN:AA006138.1; PID:g415
A:Experimental source: strain J99
C:Genetics:
A:Gene: murC
C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 76.7%; Score 33; DB 2; Length 449;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAVVNC 6
1 1 1 1 1
Db 180 CAVVNC 185

RESULT 10
T23841
hypothetical protein M88.6b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23841
R:Sulston, J.
submitted to the EMBL Data Library, June 1994
A:Reference number: Z19806
A:Accession: T23841
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-594 <WIL>
A:Cross-references: EMBL:Z34802; PIDN:CA854282.1; GSPDB:GN00021; CESP:M88.6b
A:Experimental source: clone M88
C:Genetics:
A:Gene: CESP:M88.6b
A:Map position: 3
A:Insertions: 40/3; 60/2; 101/3; 216/2; 290/1; 361/1; 411/3; 455/1; 544/3

Query Match 76.7%; Score 33; DB 2; Length 594;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVVNC 7
1 1 1 1 1
Db 488 CLAVVNC 494

RESULT 11
T18441
asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [similarity] - malaria paras
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-Jul-2002
C:Accession: T18441
R:Lawson, D.; Bowman, S.; Barrrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18441
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-610 <LAM>
A:Cross-references: EMBL:Z98547; PIDN:CA11114.1
C:Genetics:
A:Map position: 3
A:Insertions: 27/2
A>Note: C0395W
C:Superfamily: asparagine synthase (glutamine-hydrolyzing)
C:Keywords: asparagine biosynthesis; ligase
F:2-610/Product: asparagine synthase (glutamine-hydrolyzing) #status predicted <DUM>
F:2/Active site: Cys #status predicted

Query Match 76.7%; Score 33; DB 2; Length 610;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY      1 CAVPN 6
      111:11
Db      107 CAVIPN 112

RESULT 12
T23836
hypothetical protein M88.6a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23836
R:Stulson, J.
submitted to the EMBL Data Library, June 1994
A:Reference number: Z19806
A:Accession: T23836
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-610 <M1>
A:Cross-references: EMBL:Z34802; PIDN:CAA84337.1; GSPDB:GN00021; CESP:M88.6a
A:Experimental source: clone M88
C:Genetics:
A:Gene: CESP:M88.6a
A:Map position: 3
A:Insertions: 40/3; 60/2; 101/3; 216/2; 290/1; 427/3; 471/1; 560/3
Query Match
BestLocal Similarity 57.1%; Score 33; DB 2; Length 610;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 CAVPNC 7
      111:11
Db      504 CLAIPNC 510

RESULT 13
JH0579
hepatocyte growth factor precursor [validated] - human
N:Alternate names: hepatolactin A; scatter factor
C:Species: Homo sapiens (man)
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C:Accession: JH0579; J00333; A41140; B36677; A33512; A39006; PH0114; A37796; S06
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A:Title: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0579; MUID:91340155; PMID:1831432
A:Accession: JH0579
A:Molecule type: DNA
A:Residues: 1-728 <SEK>
A:Cross-references: DDBJ:D90318
A:Note: the authors translated the codon GAA for residue 662 as Gly
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
submitted to JIPID, March 1991
A:Description: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: J00333
A:Accession: J00333
A:Molecule type: DNA
A:Residues: 1-481, 'RT', '484-728 <SE2>
R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor
A:Reference number: A41140; MUID:91334393; PMID:1831266
A:Accession: A41140
A:Molecule type: mRNA
A:Residues: 1-728 <WEI>
A:Cross-references: GB:M73239; NID:9337935; PIDN:AAA64239.1; PID:9337936
R:Seiki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor
A:Reference number: A36677; MUID:91025062; PMID:2145836
A:Accession: B36677
A:Molecule type: mRNA
A:Residues: 1-728 <SR3>
A:Cross-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:9184032

```

```

A:Accession: A36677
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <SR4>
A:Cross-references: EMBL:X16323
A:Experimental source: leukocyte
R: Miyazawa, K.; Tsudouchi, H.; Naka, D.; Takahashi, K.; Ohtgaki, M.; Arakaki, N.; Nak
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
A:Reference number: A33512; MUID:89392017; PMID:2528952
A:Accession: A33512
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-728 <MIY>
A:Cross-references: GB:M29145; NID:9184041; PIDN:AAA52650.1; PID:9306846
R: Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.;
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepat
A:Reference number: A39006; MUID:91110540; PMID:1824873
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <RUB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R: Toshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayam
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native a
A:Reference number: PH0114; MUID:91207365; PMID:1826837
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43, 53-58 <YOS>
A:Experimental source: plasma
R: Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of
A:Reference number: A37796; MUID:91035621; PMID:2146276
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91, 329-344, 356-363, 'XX', 366-370, 425-434, 442-447, 'X', 449-450, 543-546, 'X
R: Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seiki, T.; Shimonishi, M.; Sugimura, A.; Ta
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: S06794; MUID:90066676; PMID:2531289
A:Accession: S06794
A:Molecule type: mRNA
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'W', 301-316, 'A', 318-335, 'K', 337-38
A:Cross-references: EMBL:X16323; NID:932081; PIDN:CAA34387.1; PID:932082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
R: Hartmann, G.; Naidini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Bir
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth f
A:Reference number: I59214; MUID:93087571; PMID:1280830
A:Accession: I59214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-288, 'ET', 'HAR>
A:Cross-references: GB:I02931; NID:9184033; PIDN:AAA52649.1; PID:9184034
R: Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth facto
A:Reference number: S15443; MUID:91200041; PMID:1826653
A:Accession: S15443
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288, 'ET', 'HAR>
A:Cross-references: EMBL:X57574; NID:932083; PIDN:CAA40802.1; PID:932084
R: Shim, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: clon
A:Reference number: I52253; MUID:92062058; PMID:1835383
A:Accession: I52253
A:Status: preliminary; translated from GB/EMBL/DBJ

```


A:Molecule type: mRNA
 A:Residues: 161-166 <SHI>
 A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997
 C:Genetics:
 A:Gene: GDB:HGR
 A:Cross-references: GDB:127524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; py
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-494/Domain: product; hepatocyte growth factor #status experimental <MAT>
 F:32-494/Domain: alpha chain #status experimental <ACH>
 F:128-206/Domain: kringle homology <KRI>
 F:211-288/Domain: kringle homology <KR2>
 F:305-383/Domain: kringle homology <KR3>
 F:391-469/Domain: kringle homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCH>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F:294,402,566,653/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:487-604/Disulfide bonds: #status predicted

Query Match 76.7%; Score 33; DB 1; Length 728;
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
 Db 377 CSQIPNC 383

RESULT 14

T06246

aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - soybea
 N:Contains: aspartate kinase (EC 2.7.2.4); homoserine dehydrogenase (EC 1.1.1.3)
 C:Species: Glycine max (soybean)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
 C:Accession: T06246
 R:Gebhardt, J.S.; Weisemann, J.M.; Matthews, B.F.
 A:Submitted to the EMBL Data Library, February 1998
 A:Description: Genes encoding the bifunctional aspartokinase-homoserine dehydrogenase fr
 A:Reference number: 215563
 A:Accession: T06246
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-909 <GEB>
 A:Cross-references: EMBL:AF049708; NID:g2970554; PIDN:AAC05983.1; PID:g2970556
 A:Experimental source: cultivar Century
 C:Genetics:
 A:Gene: AK-HSDH
 A:Introns: 47/2; 69/1; 194/1; 223/3; 325/1; 342/3; 414/2; 443/3; 481/3; 514/3; 579/3; 60
 C:Function: <AK>
 A:Description: catalyzes phosphorylation of aspartate
 C:Function: <HSD>
 A:Description: catalyzes reduction of aspartate beta-semialdehyde into homoserine
 C:Superfamily: thra bifunctional enzyme; aspartate kinase homology; homoserine dehydr
 C:Keywords: multifunctional enzyme; oxidoreductase; phosphotransferase
 F:87-552/Domain: aspartate kinase homology <DKI>
 F:555-813/Domain: homoserine dehydrogenase homology <HSD>

Query Match 76.7%; Score 33; DB 2; Length 909;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AVVPNC 7
 Db 483 AVIPNC 488

RESULT 15

T06242

aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - soy
 N:Contains: aspartate kinase (EC 2.7.2.4); homoserine dehydrogenase (EC 1.1.1.3)
 C:Species: chloroplast Glycine max (soybean)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
 C:Accession: T06242
 R:Gebhardt, J.S.; Weisemann, J.M.; Matthews, B.F.
 A:Submitted to the EMBL Data Library, February 1998
 A:Description: Genes encoding the bifunctional aspartokinase-homoserine dehydrogenase
 A:Reference number: 215563
 A:Accession: T06242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-916 <GEB>
 A:Cross-references: EMBL:AF049706; NID:g2970446; PIDN:AAC05981.1; PID:g2970447
 A:Experimental source: cultivar Century
 C:Genetics:
 A:Gene: AK-HSDH
 A:Genome: chloroplast
 C:Function: <AK>
 A:Description: catalyzes phosphorylation of aspartate
 C:Function: <HSD>
 A:Description: catalyzes reduction of aspartate beta-semialdehyde into homoserine
 C:Superfamily: thra bifunctional enzyme; aspartate kinase homology; homoserine dehydr
 C:Keywords: chloroplast; multifunctional enzyme; oxidoreductase; phosphotransferase
 F:87-552/Domain: aspartate kinase homology <DKI>
 F:555-813/Domain: homoserine dehydrogenase homology <HSD>

Query Match 76.7%; Score 33; DB 2; Length 916;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AVVPNC 7
 Db 483 AVIPNC 488

Search completed: September 4, 2003, 21:10:53
 Job time : 10.918 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11 : Search time 6.07353 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-64
Perfect score: 43
Sequence: 1 CAVVPNC 7

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	34	79.1	YAV2_XANCV	P14728 xanthomonas
2	33	76.7	SOH1_YEAST	P38633 saccharomyc
3	33	76.7	FSL3_MOUSE	Q9egc7 mus musculu
4	33	76.7	MURC_HELPJ	Q92112 helicobacte
5	33	76.7	HGE_HUMAN	P14210 homo sapien
6	32	74.4	NU6M_HYILA	Q95710 hylobates I
7	32	74.4	NU6M_PONPA	P92700 pongo pygma
8	32	74.4	Y208_MYCPN	P75486 mycoplasma
9	32	74.4	HMB1_SOYBN	P6608 glycine max
10	32	74.4	PTA_HELPJ	Q92ku4 helicobacte
11	32	74.4	PTA_MYCTU	P66254 mycobacteri
12	32	74.4	TR2M_AGRF4	P04029 agrobacteri
13	32	74.4	TR2M_AGRF4	P35017 agrobacteri
14	32	74.4	GCSR_MOUSE	P40223 mus musculu
15	31	72.1	VU5_HSVJ7	P52522 human herpe
16	31	72.1	NU6M_GORGO	Q34573 gorilla gor
17	31	72.1	PHB8_CHRYI	P45375 chromosome
18	31	72.1	Y753_STNY3	P74633 stnechocyst
19	31	72.1	Y920_TREPA	O83890 treponema p
20	31	72.1	FAT2_DROME	Q9vww1 drosophila
21	30	69.8	A63F_DROME	O46200 drosophila
22	30	69.8	TEF1_RAT	Q63467 rattus norv
23	30	69.8	PTA_CLOTM	O52593 clostridium
24	30	69.8	ARP6_CHICK	Q9dee3 gallus gall
25	30	69.8	ARP6_HUMAN	Q9gzni homo sapien
26	30	69.8	ARP6_MOUSE	Q9d864 mus musculu
27	30	69.8	ACTU_DROME	P45890 drosophila
28	30	69.8	NIFS_AZOB	P70727 azospirilla
29	30	69.8	ARGJ_STAM	O8nywt s arginine
30	30	69.8	LY41_THEMA	O9wz22 thermotoga
31	30	69.8	STE_AERPE	Q9y9n1 aeropyrum p
32	30	69.8	DL_DROME	P10041 drosophila
33	30	69.8	ARP8_YEAST	Q12386 saccharomyc

34	30	69.8	1146	1	AGEL_CAEEL	Q94125 caenorhabdi
35	29	67.4	72	1	HP11_ECTVA	P38941 ectochlorho
36	29	67.4	122	1	YEM6_YEAST	P43624 saccharomyc
37	29	67.4	170	1	YME1_THIFE	P22902 thioabacilli
38	29	67.4	175	1	PBP4_DROME	P54194 drosophila
39	29	67.4	208	1	HIS7_TRIHA	P34041 trichoderma
40	29	67.4	220	1	YHCW_BACSU	P54607 bacillus su
41	29	67.4	227	1	HIS7_MAGSR	O42621 magnaporthe
42	29	67.4	271	1	RCEL_SCHPO	O94448 schizosacch
43	29	67.4	293	1	RCE1_DROME	Q9vdc5 drosophila
44	29	67.4	302	1	WR40_ARATH	Q9san7 arabidopsis
45	29	67.4	330	1	UPAR_BOVIN	Q05588 bos taurus

ALIGNMENTS

```

RESULT 1
YAV2_XANCV          STANDARD:      PRT:      784 AA.
ID YAV2_XANCV
AC P14728:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 82 kDa avirulence protein in avrBs3 region.
OS Xanthomonas campestris (pv. vesicatoria).
OG Plasmid pXv11.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OX Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxID=341;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate 71-21;
RX MEDLINE=89384426; PubMed=2550761.
RA Bonas U., Stali R.E., Staskawicz B.;
RT Genetic and structural characterization of the avirulence gene
RL avrBs3 from Xanthomonas campestris pv. vesicatoria.";
RL Mol. Gen. Genet. 218:127-136(1989).
CC -1- DISEASE: THE TERM AVIRULENT DESCRIBES A POTENTIALLY VIRULENT
CC PATHOGEN WHICH IS UNABLE TO INDUCE DISEASE SYMPTOMS IN SPECIFIC
CC CULTIVARS OF PEPPER.
CC -1- MISCELLANEOUS: THIS IS ONE OF THE HYPOTHETICAL PROTEINS CODED BY
CC THE OPEN READING FRAMES WITHIN THE REGION REQUIRED FOR AVRBS3
CC ACTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16130;.; NOT_ANNOTATED_CDS.
CC PIR; J00317; J00317.
CC DR InterPro: IPR005042; Avirulence.
CC DR Pfam: PF03377; Avirulence; 1.
CC KW Hypothetical protein; Plasmid; Virulence.
CC SEQUENCE 784 AA; 82074 MW; 3826248411793744 CRC64;

Query Match          79.1%; Score 34; DB 1; Length 784;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVPNC 7
Db 13 CLIPNC 19

RESULT 2
SOH1_YEAST          STANDARD:      PRT:      127 AA.
ID SOH1_YEAST
AC P38633;

```

```

DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SOH1 protein.
GN SOH1 OR YGL127C OR G264.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=95073608; PubMed=7982575;
RA Fan H.-Y., Klein H.L.;
RT "Characterization of mutations that suppress the temperature-sensitive
RL growth of the hpl1 delta mutant of Saccharomyces cerevisiae.";
RN Genetics 137:945-956(1994).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96287651; PubMed=8686378;
RA Rodriguez-Belmonte E., Rodriguez Torres A.M., Tizon B., Cadahia J.L.,
RA Gonzalez-Siso I., Ramil E., Becerra M., Gonzalez-Dominguez M.,
RA Cerdan E.;
RT "Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII
RT reveals a novel member of the DnaJ family.";
RL Yeast 12:145-148(1996).
CC -I- FUNCTION: SUPPRESSES THE TEMPERATURE SENSITIVE GROWTH OF HPL1
CC MUTANT.
CC -I- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -I- SIMILARITY: Belongs to the mediator SOH1 subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: L31921; AAA35066.1; -
DR EMBL: X87252; CAA60704.1; -
DR EMBL: Z72649; CAA96836.1; -
DR PIR: S47895; S47895.
DR TRNSPAC: T03615; -
DR SCD: S0003095; SOH1.
DR GO: GO:0006281; P:DNA repair; IPT.
DR GO: GO:0006366; P:transcription from Pol II promoter; IGT.
DR KW Transcription regulation; Nuclear protein.
SQ SEQUENCE 127 AA; 14741 MW; A2PE3B62ADJC14011 CRC64;

Query Match 76.7%; Score 33; DB 1; Length 127;
Best local similarity 71.4%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Yy 1 CAVVNC 7
Db 72 CAVVNC 78

RESULT 3
FSL3_MOUSE STANDARD; PRT; 256 AA.
AC Q9EOC7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Follistatin-related protein 3 precursor (Follistatin-like 3)
DE (Follistatin-related gene protein).
GN FSL3 OR FLRG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20576231; PubMed-11010968;
RA Tsuchida K., Arai K.Y., Kuramoto Y., Yamakawa N., Hasegawa Y.,
RT Sugino H.;
RT "Identification and characterization of a novel follistatin-like
RT protein as a binding protein for the TGF-beta family.";
RT J. Biol. Chem. 275:40788-40796(2000).
RN [2]
RN RP SEQUENCE FROM N.A.
RA Uchiyama Y., Ohsawa Y., Kametaka S.;
RT "Mouse follistatin-related protein FRNG (mouse PCYF35, PC12 cell
RT derived trophic factor 35).";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=21352723; PubMed-11459787;
RA Tortorello D.V., Sidis Y., Holtzman D.A., Holmes W.E.,
RT Schneyer A.L.;
RT "Human follistatin-related protein: a structural homologue of
RT follistatin with nuclear localization.";
RL Endocrinology 142:3426-3434(2001).
CC -1- FUNCTION: Binding protein for activin and BMP-2.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: Abundantly expressed in heart, lung, kidney,
CC and testis.
CC -1- SIMILARITY: Contains 2 follistatin-like domains.
CC -1- SIMILARITY: Contains 2 Kazal-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF276238; AACG47666.1; -
DR EMBL; AB024429; BAB32663.1; -
DR EMBL; AF310616; AACG3666.1; -
DR HSSP; P37109; PCE.
DR MGD; MGI:1890391; Pstl3.
DR InterPro; IPR003645; Fcrl3.
DR InterPro; IPR002350; Kazal.
DR Pfam; PF00050; Kazal; 2.
DR SMART; SM00274; FOLN; 2.
DR SMART; SM00280; KAZAL; 2.
DR PROSITE; PS00282; KAZAL; FALSE_NEG.
RW KAZAL glycoprotein; Repeat; Signal.
KW SIGNAL.
FT CHAIN 1 23 POTENTIAL.
FT FT 24 256 FOLLISTATIN-RELATED PROTEIN 3.
FT FT 96 118 FOLLISTATIN 1.
FT FT 119 165 KAZAL-LIKE 1.
FT FT DOMAIN 168 191 KAZAL-LIKE 2.
FT FT DOMAIN 193 241 FOLLISTATIN 2.
FT FT DISULFID 119 151 KAZAL-LIKE 2.
FT FT DISULFID 123 144 BY SIMILARITY.
FT FT DISULFID 133 165 BY SIMILARITY.
FT FT DISULFID 193 227 BY SIMILARITY.
FT FT DISULFID 198 220 BY SIMILARITY.
FT FT DISULFID 209 241 BY SIMILARITY.
FT FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 256 AA; 27271 MW; 56544B8C102EE9 CRC64;

Query Match 76.7%; Score 33; DB 1; Length 256;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CAVVPC 7
I I I I I
db 117 CECVPC 123

```

```

RESULT 4
MURC_HELPJ ID MURC_HELPJ STANDARD; PRT; 449 AA.
09ZL12:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmutamate-L-alanine ligase (EC 6.3.2.8) (UDP-N-
  acetylmuramoyl-L-alanine synthetase).
GN MURC OR JHP0567.
OS Helicobacter pylori J59 (Campylobacter pylori J59).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
  Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
  Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
  Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
  Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
  Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
  Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
  gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC CC
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
  phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the murDEF family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC
DR EMBL, AE001489, AAD06138.1; -.
DR PIR, B71917, B71917.
DR HAMAP, MF_00046; -. 1.
DR InterPro: IPR000713; Mur_Ligase.
DR InterPro: IPR004101; Mur_Ligase_C.
DR InterPro: IPR005758; MurC.
DR Pfam: PF01225; Mur_Ligase_1.
DR Pfam: PF02875; Mur_Ligase_C; 1.
DR TIGRFAMs: TIGR01082; murC; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
  ATP-binding; Complete proteome.
FT NP-BIND 121 127 ATP (POTENTIAL).
SQ SEQUENCE 449 AA; 50906 MW; 4EDDCCF8AF5D74F CRC64;
Query Match 76.7%; Score 33; DB 1; Length 449;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAVPN 6
Db 180 CAIVPN 185

```

```

DE (Hepatopocetin A).
GN HGF OR HPTA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
  Seki T., Hagiya M., Shimomishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene."
  Gene 102:213-219(1991).
RL
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9392017; PubMed=2528952;
  Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
  Araiaki N., Nakayama H., Hiroo S., Sakiyama O., Takahashi K.,
  Gohda E., Daikuhara Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
  growth factor."
  Biochem. Biophys. Res. Commun. 163:967-973(1989).
RL
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Leukocyte;
  MEDLINE=91025062; PubMed=2145836;
  Seki T., Ihara I., Sugimura A., Shimomishi M., Nishizawa T.,
  Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT "Isolation and expression of cDNA for different forms of hepatocyte
  growth factor from human leukocyte."
  Biochem. Biophys. Res. Commun. 172:321-327(1990).
RL
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RX TISSUE=Liver;
  MEDLINE=9006676; PubMed=2531289;
  Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimomishi M.,
  Sugimura A., Tashiro K., Shimizu S.;
RT "Molecular cloning and expression of human hepatocyte growth factor."
  Nature 342:440-443(1989).
RL
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryonic fibroblast;
  MEDLINE=91334393; PubMed=1831266;
  Weidner K.M., Araiaki N., Hartmann G., Vandekerckhove J., Weingart S.,
  Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
  Birnmeier W.;
RT "Evidence for the identity of human scatter factor and human
  hepatocyte growth factor."
  Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
RL
RN [6]
RP SEQUENCE FROM N.A.
RX Courtney L., Elliot G., Angell S.;
  Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 249-695 FROM N.A.
RX MEDLINE=91369928; PubMed=1832556;
  Miyazawa K., Kitamura A., Kitamura N.;
RT "Structural organization and the transcription initiation site of the
  human hepatocyte growth factor gene."
  Biochemistry 30:9170-9176(1991).
RL
RN [8]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=91207365; PubMed=1826837;
  Yoshizawa Y., Araiaki N., Naka D., Takahashi K., Hiroo S., Kondo J.,
  Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
  Hishida T., Daikuhara Y.;
RT "Identification of the N-terminal residue of the heavy chain of both
  native and recombinant human hepatocyte growth factor."
  Biochem. Biophys. Res. Commun. 175:660-667(1991).
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-476.
RX MEDLINE=93129192; PubMed=1482346;
  Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,

```

RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 RT on the alpha chain."
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [10]
 RP MUTAGENESIS.
 RA MEDLINE-92331602; PubMed1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 RT identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding."
 RL EMBO J. 11:2503-2510(1992).
 RN [11]
 RP STRUCTURE BY NMR OF 31-127.
 RA MEDLINE-98154323; PubMed-9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wangfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site."
 RL Structure 6:109-116(1998).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RA MEDLINE-99036858; PubMed-9817840;
 RA Ultsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 RT factor at 2.0-A resolution."
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS. SEEMS TO BE AN HEPATOPROTECTIVE FACTOR. AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90334; BAA14348.1; -
 DR EMBL; D90318; BAA14348.1; JOINED.
 DR EMBL; D90319; BAA14348.1; JOINED.
 DR EMBL; D90320; BAA14348.1; JOINED.
 DR EMBL; D90322; BAA14348.1; JOINED.
 DR EMBL; D90323; BAA14348.1; JOINED.
 DR EMBL; D90324; BAA14348.1; JOINED.
 DR EMBL; D90325; BAA14348.1; JOINED.
 DR EMBL; D90326; BAA14348.1; JOINED.
 DR EMBL; D90327; BAA14348.1; JOINED.
 DR EMBL; D90328; BAA14348.1; JOINED.
 DR EMBL; D90329; BAA14348.1; JOINED.
 DR EMBL; D90330; BAA14348.1; JOINED.
 DR EMBL; D90331; BAA14348.1; JOINED.
 DR EMBL; D90332; BAA14348.1; JOINED.
 DR EMBL; D90333; BAA14348.1; JOINED.
 DR EMBL; M29145; AAA52650.1; -
 DR EMBL; M60718; AAA52648.1; -
 DR EMBL; X16323; CAA34387.1; -
 DR EMBL; M73239; CAA34387.1; -
 DR EMBL; M73240; AAA64297.1; -
 DR EMBL; AC004960; AAC71655.1; -
 DR EMBL; M75983; AAG53460.1; -
 DR EMBL; M75972; AAG53460.1; JOINED.
 DR EMBL; M75973; AAG53460.1; JOINED.
 DR EMBL; M75974; AAG53460.1; JOINED.
 DR EMBL; M75975; AAG53460.1; JOINED.
 DR EMBL; M75976; AAG53460.1; JOINED.

DR EMBL; M75977; AAG53460.1; JOINED.
 DR EMBL; M75978; AAG53460.1; JOINED.
 DR EMBL; M75979; AAG53460.1; JOINED.
 DR EMBL; M75980; AAG53460.1; JOINED.
 DR EMBL; M75981; AAG53460.1; JOINED.
 DR EMBL; M75982; AAG53460.1; JOINED.
 DR PIR; JH0579; JH0579.
 DR PDB; 2HGF; 24-JUN-98.
 DR PDB; 1BHT; 18-NOV-98.
 DR PDB; 1NK1; 29-DEC-99.
 DR MEROPS; S01.976; -.
 DR GlycoSuiteDB; P14210; -.
 DR Genew; HGNC:4893; HGF.
 DR MIM; 142409; -.
 DR GO; GO:0008083; F: growth factor activity; NAS.
 DR GO; GO:0007067; P: mitosis; NAS.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00024; Kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_Spec; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;
 KW Signal; 3d-structure; Polymorphism; Pyroliidone carboxylic acid.
 FT CHAIN 1 31
 FT CHAIN 495 728 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
 FT MOD_RES 32 32 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
 FT DOMAIN 32 127 PYROLIDONE CARBOXYLIC ACID.
 FT DOMAIN 128 206 PAP.
 FT DOMAIN 211 288 KRINGLE 1.
 FT DOMAIN 305 383 KRINGLE 2.
 FT DOMAIN 391 469 KRINGLE 3.
 FT DOMAIN 495 728 KRINGLE 4.
 FT DISULFID 70 96 SERINE PROTEASE-LIKE.
 FT DISULFID 74 84
 FT DISULFID 128 206
 Query Match 76.7%; Score 33; DB 1; Length 728;
 Best Local Similarity 57.1%; Pred. No. 36;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1-CAVPPNC 7
 Db 377 CSQIPNC 383
 RESULT 6
 NUGM_HYLLA
 ID NUGM_HYLLA STANDARD; PRT; 174 AA.
 AC 095710;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 6 (BC 1.6.5.3).
 GN MTN6 OR ND6 OR NADH6.
 OS Hylobates lar (Common gibbon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OC NCBI_Taxid=9580;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-Ester;
RA Arnason U., Gullberg A., Xu X.;
RT "A complete mitochondrial DNA molecule of the white-handed gibbon,
RT Hylobates lar, and comparison among individual mitochondrial genes of
RT all hominoid genera.";
RL Hereditas 124:185-189(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99256; CAAG7639.1; -.
DR PIR: T11844; T11844.
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18556 MW; A956501979D20F03 CRC64;

Query Match
Best Local Similarity 74.4%; Score 32; DB 1; Length 174;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAVVNC 7
Db 40 CAVILNC 46

RESULT 7
NU6M_PONPA STANDARD; PRT; 174 AA.
AC P92700;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
GN MTND6 OR ND6 OR NADH6.
OS Pongo pygmaeus abelli (Sumatran orangutan).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISolate YN93-312;
RX MEDLINE=97032590; PubMed=8875856;
RA Xu X., Arnason U.;
RT "The mitochondrial DNA molecule of Sumatran orangutan and a molecular
RT proposal for two (Bornean and Sumatran) species of orangutan.";
RL J. Mol. Evol. 43:431-437(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97707; CAAG6294.1; -.
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18562 MW; A3BE85628CB18F8A CRC64;

Query Match
Best Local Similarity 74.4%; Score 32; DB 1; Length 174;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAVVNC 7
Db 40 CAVILNC 46

RESULT 8
Y208_MYCPN STANDARD; PRT; 196 AA.
ID Y208_MYCPN
AC P75486;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG208 homolog (H10_orf196).
GN MN291 OR MP544.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000053; AAB96192.1; -.
DR PIR: S73870; S73870.
KM Hypothetical protein: Complete proteome.
SQ SEQUENCE 196 AA; 22464 MW; E8E55512BC381BF CRC64;

Query Match
Best Local Similarity 74.4%; Score 32; DB 1; Length 196;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAVVNC 7
Db 90 CTVNPN 96

RESULT 9
HMBL_SOYBN STANDARD; PRT; 379 AA.
ID HMBL_SOYBN
AC P46608;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein SBH1.
GN H1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Somatic embryo;
RX MEDLINE=94169300; PubMed=7907232;
RA Ma H., McMullen M.D., Finer J.J.;
RT "Identification of a homeobox-containing gene with enhanced
RT expression during soybean (Glycine max L.) somatic embryo
RT development.";

```

RL Plant Mol. Biol. 24:465-473(1994).
 CC 1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR INVOLVED IN EARLY
 CC EMBRYONIC DEVELOPMENT. PROBABLY BINDS TO THE DNA SEQUENCE 5'-TGAC-
 CC 3'.
 CC 1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC 1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN EMBRYONIC TISSUES. WEAKLY
 CC DETECTED IN STEMS AND HYPOCOTYL.
 CC 1- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN THE EMBRYO PROLIFERATION
 CC STAGE, INCREASES DURING EARLY SOMATIC EMBRYO DEVELOPMENT AND
 CC DECREASES THEREAFTER.
 CC 1- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEOBOX FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L13663; AAA20882.1; -
 DR PIR: S42543; S42543.
 DR HSSP: P41778; ID06.
 DR TRANSFAC: T04052; -
 DR InterPro: IPR005539; ELK.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR005340; KNOX1.
 DR InterPro: IPR005541; KNOX2.
 DR Pfam: PF03789; ELK; 1.
 DR Pfam: PF00046; homeobox; 1.
 DR Pfam: PF03790; KNOX1; 1.
 DR Pfam: PF03791; KNOX2; 1.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Transcription regulation;
 KW Activator.
 FT DOMAIN 102 107 POLY-HIS.
 FT DOMAIN 109 113 POLY-ASN.
 FT DOMAIN 115 122 POLY-SER.
 FT DOMAIN 244 248 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 259 282 ELK DOMAIN.
 FT DNA_BIND 283 345 HOMEBOX (TALE-TYPE).
 SQ SEQUENCE 379 AA; 42374 MW; 886B020029EDD5E4 CRC64;
 QY Query Match 74.4%; Score 32; DB 1; Length 379;
 Best Local Similarity 57.1%; Pred. No. 31;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 CAVPNC 7
 DB 56 CLIPNC 62
 RESULT 10
 PTA_HELPJ STANDARD: PRT; 519 AA.
 AC Q9ZK04;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
 GN PTA OR JH0841.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carnell G.,

RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC 1- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate = CoA + acetyl
 CC phosphate.
 CC 1- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
 CC ACETYL-COA.
 CC 1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC 1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE PHOSPHATE
 CC ACETYLTRANSFERASE AND BUTYRYLTRANSFERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE001513; AAD06419.1; -
 DR PIR: D71881; D71881.
 DR InterPro: IPR004614; Pta.
 DR InterPro: IPR002505; PTA_PTB.
 DR Pfam: PF01515; PTA_PTB; 1.
 DR TIGRFAMS: TIGR00651; pta; 1.
 KW Transferase; Acyltransferase; Complete proteome.
 FT DOMAIN 196 519 PHOSPHATE ACETYLTTRANSFERASE.
 SQ SEQUENCE 519 AA; 56737 MW; 3E5A71FCBBBD37 CRC64;
 QY Query Match 74.4%; Score 32; DB 1; Length 519;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CAVPN 6
 DB 361 CALIPN 366
 RESULT 11
 PTA_MYCTU STANDARD: PRT; 690 AA.
 AC P96254;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
 GN PTA OR RV0408 OR MT0421 OR MYCY22610.04.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Jags K., Kiroh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.,
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate -> CoA + acetyl
CC phosphate.
CC -1- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
CC ACETYL-COA.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; CONTAINS A DOMAIN WITH SOME
CC SIMILARITIES WITH COB/BHD.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE PHOSPHATE
CC ACETYLTRANSFERASE AND BUTYRYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z84724; CAB06578.1; -
DR EMBL: AE006946; AAK44645.1; -
DR PIR: F70628; F70628.
DR TIGR: MT0421; -
DR Tuberculist; RV0408; -
DR InterPro: IPR004614; Pla.
DR InterPro: IPR002505; PTA_PTB.
DR Pfam: PF01515; PTA_PTB; 1.
DR TIGRFAMs: TIGR00651; pta; 1.
KW Transferase; Acyltransferase; Complete proteome.
FT DOMAIN 365 690 PHOSPHATE ACETYLTRANSFERASE.
SQ SEQUENCE 690 AA; 72948 MW; C01CA12AF2810CCE CRC64;

Query Match 74.4%; Score 32; DB 1; Length 690;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAVVNP 6
DB 530 CALIPN 535

RESULT 12
ID TR2M_AGR74 STANDARD: PRT; 755 AA.
AC P04029;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN TMS1.
OS Agrobacterium tumefaciens (strain Ach5), and
OS Agrobacterium tumefaciens.
OC Plasmid pTiAch5, and Plasmid pTiA6NC.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176298, 358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Ach5; PLASMID-pTiAch5;
RX MEDLINE=84207942; Pubmed=6327292;
RA Gielen J., de Beuckeleer M., Seurinck J., Deboeck F., de Greve H.,
RA Lemmens M., van Montagu M., Schell J.,
RT "The complete nucleotide sequence of the TL-DNA of the Agrobacterium
RT tumefaciens plasmid pTiAch5."
RL EMBO J. 3:835-846(1984).
RN [2]
RP SEQUENCE FROM N.A.

RC PLASMID-pTiA6NC;
RX MEDLINE=84170374; Pubmed=6584906;
RA Klee H., Montoya A., Horodyski F., Lichtenstein C., Garfinkel D.,
RA Fuller S., Flores C., Peschon J., Nester E., Gordon M.,
RT "Nucleotide sequence of the tms genes of the pTiA6NC octopline Ti
RT plasmid: two gene products involved in plant tumorigenesis."
RL Proc. Natl. Acad. Sci. U.S.A. 81:1728-1732(1984).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) -> indole-3-acetamide +
CC CO(2) + H(2)O.
CC -1- PATHWAY: Biosynthesis of auxins from tryptophan, first step.
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF ACH5.
CC -1- SIMILARITY: SIGNIFICANT HOMOLOGY TO THE ADENINE BINDING REGION OF
CC P-HYDROXYBENZAZONE HYDROXYLASE FROM P. FLUORESCENS. IT SEEMS THAT
CC THIS PROTEIN BINDS ADENINE EITHER AS SUBSTRATE OR COFACTOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02554; AAA92550.1; -
DR PIR: A04497; QOAG4T.
DR InterPro: IPR000759; Adnrx_reductase.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR006064; Glycosidase.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
DR Pfam: PF02027; RoLB_RoIC; 1.
DR PRINTS: PR00419; ADXRDYASE.
KW Oxidoreductase; Monooxygenase; Auxin biosynthesis; Crown gall tumor;
FT VARIANT 718 719 NR -> IQ (IN PTIA6NC).
FT VARIANT 721 721 P -> A (IN PTIA6NC).
SQ SEQUENCE 755 AA; 83947 MW; 9FD2B83FEA001AAD CRC64;

Query Match 74.4%; Score 32; DB 1; Length 755;
Best Local Similarity 57.1%; Pred. No. 59;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNP 7
DB 726 CALIHC 732

RESULT 13
ID TR2N_AGRV1 STANDARD: PRT; 755 AA.
AC P25017;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN IAAM.
OS Agrobacterium vitis (Rhizobium vitis).
OC Plasmid pTiV4.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=373;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TM4;
RX MEDLINE=91329707; Pubmed=1868204;
RA Bonnard G., Vincent F., Otten L.,
RT "Sequence of Agrobacterium tumefaciens biotype III auxin genes."
RT Plant Mol. Biol. 16:733-738(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CG474;
RA Otten L., de Ruffray P.,
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

```

CC -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) = indole-3-acetamide +
CC Co(2) + H(2)O.
CC -1- PATHWAY: Biosynthesis of auxins from tryptophan; first step.
CC -1- SIMILARITY: SIGNIFICANT HOMOLOGY TO THE ADENINE BINDING REGION OF
CC P-HYDROXYBENZOATE HYDROXYLASE FROM P. FLUORESCENS. IT SEEMS THAT
CC THIS PROTEIN BINDS ADENINE EITHER AS SUBSTRATE OR COFACTOR.
CC -1- SIMILARITY: STRONG, WITH ITS COUNTERPARTS FROM A. TUMEFACIENS
CC PLASMIDS PT115955, PT1ACH5 AND PT1A6NC.
CC -1- CAUTION: THE PLASMID PT11A4 CARRIES TWO T-REGIONS, THE TA AND
CC TB REGION, BOTH OF WHICH HAVE A FUNCTIONAL TAAM GENE, WITH LOW
CC HOMOLOGY BETWEEN THEM.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL: X56185; CAA39646.1; -.
CC EMBL: U83987; ABA41874.1; -.
CC InterPro: IPR000759; Adrndx_reductase.
CC InterPro: IPR002937; Amino_oxidase.
CC InterPro: IPR006064; Glycosidase.
CC InterPro: IPR000205; NAD_binding.
CC Pfam: PF01593; Amino_oxidase; 1.
CC Pfam: PF02027; Ro1B_Ro1C; 1.
CC PRINTS: PR00419; ADKRDYASE.
CC OXfoldordatabase: Monooxygenase; Auxin biosynthesis; Crown gall tumor;
CC T1DNA; Plasmid.
CC KW SEQUENCE 755 AA; 83972 MW; 6FA63E502343136F CRC64;
CC SQ
OY Query Match 74.4%; Score 32; DB 1; Length 755;
DB Best/Local Similarity 57.1%; Pred No. 59;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 CAVVENC 7
DB 726 CALHNC 732
RESULT 14
GCSR_MOUSE STANDARD; PRT; 837 AA.
AC P40223;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Granulocyte colony stimulating factor receptor precursor (G-CSF-R).
GN CSF3R OR CSFGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90235283; PubMed=2158861;
RA Fukunaga R., Ishizaka-Ikeda E., Seto Y., Nagata S.;
RT "Expression cloning of a receptor for murine granulocyte colony-
RT stimulating factor.";
RL Cell 61:341-350(1990).
RN [2]
RP STRUCTURE BY NMR OF 225-333.
RA MEDLINE=9731327; PubMed=9187659;
RA Yamasaki K., Naito S., Anaguchi H., Ohkubo T., Ota Y.;
RT "Solution structure of an extracellular domain containing the MSXMS
RT motif of the granulocyte colony-stimulating factor receptor and its
RT interaction with ligand.";
RL Nat. Struct. Biol. 4:498-504(1997).
CC -1- FUNCTION: RECEPTOR FOR GRANULOCYTE COLONY-STIMULATING FACTOR. IN
CC ADDITION IT MAY FUNCTION IN SOME ADHESION OR RECOGNITION EVENTS AT
CC THE CELL SURFACE.

```

CC	-1	SUBUNIT: DIMER (PROBABLE).		
CC	-1	SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1	TISSUE SPECIFICITY: FOUND IN BONE MARROW.		
CC	-1	SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.		
CC	-1	SIMILARITY: Contains 5 fibronectin type III domains.		
CC	-1	SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.		
CC		-----		
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC		the European Bioinformatics Institute. There are no restrictions on its		
CC		use by non-profit institutions as long as its content is in no way		
CC		modified and this statement is not removed. Usage by and for commercial		
CC		entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC		or send an email to license@isb-sib.ch).		
CC		-----		
DR	EMBL:	M58288: AAA37673.1; -.		
DR	PIR:	A34898; A34898.		
DR	PDB:	1GCF; 22-OCT-97.		
DR	PDB:	1GCT; 22-OCT-97.		
DR	PDB:	1CD9; 08-MAR-00.		
DR	PDB:	1PGR; 08-MAR-00.		
DR	MGD:	MGI:1339755; Csf3r.		
DR	GO:	GO:0030593; P:neutrophil chemotaxis; IMP.		
DR	InterPro:	IPR002896; CRIA.		
DR	InterPro:	IPR003961; FN_III.		
DR	InterPro:	IPR003529; Hemtopoptn_L_F2.		
DR	SMART:	PF00041; Fn3; 3.		
DR	SMART:	SMO0060; FN3; 2.		
DR	PROSITE:	PS01353; HEMATOPO_REC_L_F2; 1.		
KW		Cell adhesion; Receptor; Repept; Signal; Transmembrane;		
KW		Immunoglobulin domain; Glycoprotein; 3d-structure.		
FT	SIGNAL	1	POTENTIAL.	
FT	CHAIN	26	GRANULOCYTE COLONY STIMULATING FACTOR	
FT			RECEPTOR.	
FT	DOMAIN	26	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	627	POTENTIAL.	
FT		651	CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN	26	IG-LIKE C2-TYPE.	
FT	DOMAIN	122	FIBRONECTIN TYPE-III 1.	
FT	DOMAIN	229	FIBRONECTIN TYPE-III 2.	
FT	DOMAIN	334	FIBRONECTIN TYPE-III 3.	
FT	DOMAIN	432	FIBRONECTIN TYPE-III 4.	
FT	DOMAIN	529	FIBRONECTIN TYPE-III 5.	
FT	DISULFID	132	BY SIMILARITY.	
FT	DISULFID	249	BY SIMILARITY.	
FT	DISULFID	267	BY SIMILARITY.	
FT	CARBOHYD	51	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	94	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	129	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	186	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	279	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	352	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	408	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	474	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	487	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	582	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	613	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	STRAND	222	235	
FT	STRAND	249	255	
FT	STRAND	266	274	
FT	STRAND	281	286	
FT	STRAND	290	296	
FT	STRAND	303	311	
FT	STRAND	323	328	
SO	SEQUENCE	837 AA:	93406 MW:	422955E989E2C8531 CRC64:
QY	Query Match	74.4%;	Score 32;	DB 1; Length 837;
	Best Local Similarity	57.1%;	Pred. No. 65;	
	Matches	4; Conservative	1; Mismatches	2; Indels
DB		46 CITSFNC 52		0;

RESULT 15

```

ID VU5_HSV7J STANDARD: PRT: 865 AA.
AC P52522;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE U5 protein.
GN U5.
OS Human herpesvirus (type 7 / strain JI) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-6 AND
CC HSV-7 U4, HSV-6 AND HSV-7 U5 AND HCMV UL27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43400; AAC54669.1; -
DR PIR; T41909; T41909.
SQ SEQUENCE 865 AA; 100187 MW; 3D3F4770B0FFACT7 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 865;
Best Local Similarity 71.4%; Pred. NO. 68;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CAVVNC 7
| | | |
Db 199 COVERNC 205

```

Search completed: September 4, 2003, 21:01:47
 Job time : 6.07353 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47 ; Search time 28.7206 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-64

Perfect score: 43

Sequence: 1 CAVVPMC 7

Scoring table: BLOSUM62

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPREMBL_23:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_phage:*
- 11: sp_plant:*
- 12: sp_rodent:*
- 13: sp_virus:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	90.7	1704	5 094446	094446 chironomus
2	37	86.0	100	6 095KE3	095KE3 macaca fasc
3	37	86.0	409	17 090XJ6	090XJ6 sulfolobus
4	35	81.4	169	10 091G02	091G02 oryza sativ
5	34	79.1	24	10 08W1W1	08W1W1 zea mays (m
6	34	79.1	126	11 08BJU8	08BJU8 mus musculu
7	34	79.1	212	4 09P096	09P096 homo sapien
8	34	79.1	312	10 091J09	091J09 oryza sativ
9	34	79.1	316	4 08NCX1	08NCX1 homo sapien
10	34	79.1	368	16 053303	053303 mycobacteri
11	34	79.1	375	16 08YBN2	08YBN2 bruceella me
12	34	79.1	375	16 08FWP5	08FWP5 bruceella su
13	34	79.1	415	16 098AG8	098AG8 rhizobium l
14	34	79.1	433	12 091L91	091L91 white spot
15	34	79.1	446	12 08N5V5	08N5V5 homo sapien
16	34	79.1	500	11 08BLL7	08BLL7 mus musculu

17	34	79.1	558	4 08NFR0	08NFR0 homo sapien
18	34	79.1	586	4 08NFR1	08NFR1 homo sapien
19	34	79.1	605	4 08NFR2	08NFR2 homo sapien
20	34	79.1	605	11 08CG08	08CG08 mus musculu
21	34	79.1	644	5 061834	061834 caenorhabdi
22	34	79.1	2592	5 0810L2	0810L2 metapneus
23	33	76.7	73	6 08N052	08N052 macaca fasc
24	33	76.7	223	4 08W74	08W74 homo sapien
25	33	76.7	327	4 096S96	096S96 homo sapien
26	33	76.7	321	4 096S27	096S27 homo sapien
27	33	76.7	400	16 053697	053697 mycobacteri
28	33	76.7	400	16 08VKR4	08VKR4 mycobacteri
29	33	76.7	553	13 08UUP5	08UUP5 xenopus lae
30	33	76.7	594	5 09U3A0	09U3A0 caenorhabdi
31	33	76.7	610	5 021604	021604 caenorhabdi
32	33	76.7	610	5 077330	077330 plasmodium
33	33	76.7	751	10 08LKR4	08LKR4 arabidopsis
34	33	76.7	752	4 013597	013597 homo sapien
35	33	76.7	893	4 08N6Z0	08N6Z0 homo sapien
36	33	76.7	909	10 065027	065027 glycine max
37	33	76.7	916	8 063067	063067 glycine max
38	33	76.7	931	5 09SP09	09SP09 tetrahyma
39	33	76.7	1209	10 09SGS6	09SGS6 arabidopsis
40	33	76.7	1550	10 091G11	091G11 arabidopsis
41	33	76.7	1698	5 094438	094438 chironomus
42	33	76.7	1891	16 08EZS3	08EZS3 leptospira
43	32	74.4	130	4 08TAX3	08TAX3 homo sapien
44	32	74.4	147	4 09P0T3	09P0T3 homo sapien
45	32	74.4	157	2 09R467	09R467 agrobacteri

ALIGNMENTS

RESULT 1

Q94446 PRELIMINARY; PRT; 1704 AA.

AC Q94446: 01-FEB-1997 (TREMBLrel. 02, Created)

DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)

DE 220 kDa silk protein.

GN SP220.

OS Chironomus thummi (midge).

CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae; Chironomidae; Chironominae; Chironomus.

OX NCBI_TaxID=7154;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Salivary gland;

RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;

RT "Extraordinary conservation of cysteines among homologous Chironomus silk proteins sp185 and sp220."

RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL: U54641; AAA99804.1; -

DR InterPro: IPR004153; CXXC_repeat.

DR InterPro: IPR006209; EGF_like.

DR Pfam: PF0128; CXXC; 69.

DR PROSITE: PS00022; EGF_1; 2.

DR PROSITE: PS01186; EGF_2; 1.

SQ SEQUENCE 1704 AA; 185746 MW; 3A3F20247C8F1E28 CRC64;

Query Match 90.7%; Score 39; DB 5; Length 1704;
Best Local Similarity 85.7%; Pred. NO. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVPMC 7
DB 1442 CATVNC 1448

RESULT 2

095KE3
ID 095KE3 PRELIMINARY: PRT: 100 AA.
AC 095KE3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
OS Hypothetical 11.5 kDa protein.
OC Macaca fascicularis (Grab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
SU Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062929; BAB60725.1; -
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 11512 MW; 1AB73A899B73FEA8 CRC64;

Query Match 86.0%; Score 37; DB 6; Length 100;
Best Local Similarity 71.4%; Pred. No. 3.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
ID 1:1111
DB 36 CTVAPNC 42

RESULT 3
O9UXJ6 PRELIMINARY: PRT: 409 AA.
AC O9UXJ6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE UDP-glucose dehydrogenase (UDP-glucose 6-dehydrogenase) (UGD)
(EC 1.1.1.22).
GN UGD OR SS00810 OR ORF-C39_020.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Avez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Helcamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL PLoS. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; Y18930; CAB57493.1; -
DR EMBL; AB06704; AAK41109.1; -
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR006109; NAD_gly3p_dom.
DR InterPro: IPR001732; UDPG_MGDP_dh.
DR Pfam: PF00984; UDPG_MGDP_dh; 1.
DR Pfam: PF03720; UDPG_MGDP_dh; 1.
DR Pfam: PF03721; UDPG_MGDP_dh; 1.
DR Pfam: PF03721; UDPG_MGDP_dh; 1.
DR ProDom: PD001278; NAD_gly3p_dom; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 409 AA; 44837 MW; 22EAB1AB796809D6 CRC64;

Query Match 86.0%; Score 37; DB 17; Length 409;

Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVPNC 7
ID 1:1111
DB 213 CTVAPNC 219

RESULT 4
O91G02 PRELIMINARY: PRT: 169 AA.
AC O91G02;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ESTS AU071032(R10630).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone: P0510F03."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002486; BAB03366.1; -
DR Gramene; O91G02; -
DR InterPro: IPR003657; WRKY.
DR Pfam: PF03106; WRKY; 1.
DR PROSITE: PS50811; WRKY; 1.
SQ SEQUENCE 169 AA; 18275 MW; D41B876772C0A302 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 169;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAVPNC 7
ID 1:1111
DB 135 CTVAPNC 141

RESULT 5
O8W1W1 PRELIMINARY: PRT: 24 AA.
AC O8W1W1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 2.7 kDa protein (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD Clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Ananiev E.V., Lorentzen J., Bruggemann E.;
RT "Microsatellite megatracts in the maize (Zea mays L.) genome."
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413055; AAL55292.1; -
DR InterPro: IPR001594; Znf_DHHC.
DR PROSITE: PS50216; ZF_DHHC; 1.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 24
FT NON_TER 24
SQ SEQUENCE 24 AA; 2742 MW; 3F491B8A538ADF33 CRC64;

Query Match 79.1%; Score 34; DB 10; Length 24;
Best Local Similarity 57.1%; Pred. No. 3.6;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
1:111
DB 10 CSICPNC 16

RESULT 6

Q8BJU8 PRELIMINARY; PRT: 126 AA.
ID Q8BJU8
AC Q8BJU8
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Weekly similar to sodium/potassium/calcium exchanger 3 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK079049; BAC37514.1;
SO SEQUENCE 126 AA; 14052 MW; 5F659AD717CCCD89 CRC64;

Query Match 79.1%; Score 34; DB 11; Length 126;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVNC 7
1:111
DB 35 CVTIPNC 41

RESULT 7

Q9P096 PRELIMINARY; PRT: 212 AA.
ID Q9P096
AC Q9P096;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE HSPC285 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang O.H., Chen S.J., Chen Z.;
RT "Human partial CDS from cd34+ stem cells."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF1611403; AAF28963.1;
FT NON_TER
SO SEQUENCE 212 AA; 22078 MW; 12E1194E7A8F75D1 CRC64;

Query Match 79.1%; Score 34; DB 4; Length 212;
Best Local Similarity 71.4%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVVNC 7
1:111
DB 22 CAELPNC 28

RESULT 8

Q9LJ09

ID Q9LJ09 PRELIMINARY; PRT: 312 AA.

AC Q9LJ09;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to actin-like protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:p0667A10."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBICUOUSLY EXPRESSED
CC IN ALL EUKARYOTIC CELLS (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE ACTIN FAMILY.
DR EMBL: AP001073; BAA89581.1; -
DR Gramene: Q9LJ09; -
DR InterPro: IPR004000; Actin-like.
DR Pfam: PF00022; actin; 1.
DR SMART: SM00268; ACTIN; 1.
DR Structural protein.
SO SEQUENCE 312 AA; 34639 MW; EB7531627A3C5A09 CRC64;

Query Match 79.1%; Score 34; DB 10; Length 312;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVEVNC 7
111111
DB 29 AVEVNC 34

RESULT 9

Q8NCX1 PRELIMINARY; PRT: 316 AA.
ID Q8NCX1
AC Q8NCX1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN DKF2P434P185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Duesterhoeft A., Lauber J., Mewes H.W., Well B., Wiemann S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL834225; CAD38903.1; -
KW Hypothetical protein.
SO SEQUENCE 316 AA; 35263 MW; 1FA9E29F5A6DFD35 CRC64;

Query Match 79.1%; Score 34; DB 4; Length 316;
Best Local Similarity 57.1%; Pred. No. 40;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVVNC 7
1:111
DB 132 CVTIPNC 138

RESULT 10

Q53303

PRELIMINARY; PRT: 368 AA.

AC 053303;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative alcohol dehydrogenase (Zinc-binding dehydrogenase).
 GN ADHD OR RV3086 OR MV013.07 OR MT3171.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterinae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295967; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Baahm D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX Fieischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC 1- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY (BY SIMILARITY).
 CC 1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.

CC EMBL: AL021309; CA16144.1; -
 DR EMBL: AE007134; AAK47507.1; -
 DR HSSP: PA0394; IAGN.
 DR TIGR: MT3171; -
 DR Tuberculist: RV3086; -
 DR InterPro: IPR002328; Adh_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR000515; BPD_transp.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; Adh_ZINC; 1.
 DR PROSITE: PS00402; BPD_TRANSF_TNN_MEMBER; 1.
 DR Oxidoreductase; Zinc; Complete proteome.
 KW SEQUENCE 368 AA; 38341 MW; A03FCB6F3C8003AF CRC64;
 SQ

Query Match 79.1%; Score 34; DB 16; Length 368;
 Best Local Similarity 57.1%; Pred. No. 46;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAVPNC 7
 I : I I I I
 DR 85 CSFIPNC 91

RESULT 11

OBEBN2 PRELIMINARY; PRT; 375 AA.
 AC 08YBN2;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alcohol dehydrogenase (EC 1.1.1.1).
 GN BMEI10867.
 OS Brucella melitensis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kaprali V., Redkar R.J., Patra G., Muijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Rezak G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Telleson J.-J.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009720; AL54109.1; -
 DR InterPro: IPR002328; Adh_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; Adh_ZINC; 1.
 DR Oxidoreductase; Complete proteome.
 KW SEQUENCE 375 AA; 39567 MW; D8F4009E96CDBDEA CRC64;
 SQ

Query Match 79.1%; Score 34; DB 16; Length 375;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVPNC 7
 I : I I I I
 DB 93 CVFVPC 99

RESULT 12

OBFBP5 PRELIMINARY; PRT; 375 AA.
 AC 08FBP5;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE Alcohol dehydrogenase, zinc-containing.
 GN BRA0401.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Rietveld H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL: AE014539; AAN33598.1; -
 DR TIGR: BRA0401; -
 KW Complete proteome.
 SQ SEQUENCE 375 AA; 39582 MW; 779A71D041C367AA CRC64;
 SQ

Query Match 79.1%; Score 34; DB 16; Length 375;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAVPNC 7
 I : I I I I
 DB 93 CVFVPC 99


```

RESULT 13
O98AG8 PRELIMINARY: PRT: 415 AA.
ID O98AG8:
AC O98AG8:
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE 8-amino-7-oxononanoate synthase.
GN ML6006.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003008; BAB52362.1; -
DR InterPro: IPR004839; AminoTransferase/2.
DR Pfam: PF00155; aminotran_1.2; 1.
KW Complete proteome.
SQ SEQUENCE 415 AA; 44108 MW; EEE219D749AE997 CRC64;

Query Match
Best Local Similarity 79.1%; Score 34; DB 16; Length 415;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
DB 339 CGVVPSC 345

RESULT 14
O91L91 PRELIMINARY: PRT: 433 AA.
ID O91L91:
AC O91L91:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE ORF161 (MSV338) (WSSV394).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21342572; PubMed=11448154;
RX van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.R., Vlak J.M.;
RT "The white spot syndrome virus DNA genome sequence."
RL Virology 286:7-22(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.R., Vlak J.M.;
RT Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=21548311; PubMed=11689662;
RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus."
RL J. Virol. 75:11811-11820(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

```

```

RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Taiwan;
RX MEDLINE=20517548; PubMed=11062040;
RA Tsal M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase."
RL Virology 277:100-110(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Taiwan;
RX MEDLINE=21844071; PubMed=11853398;
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells."
RL Virology 293:44-53(2002).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Taiwan;
RA Lo C.-F., Kou G.-H.;
RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF369029; AAK77830.1; -
DR EMBL; AF332093; AAL33340.1; -
DR EMBL; AF440570; AAL89262.1; -
SQ SEQUENCE 433 AA; 48217 MW; 387FDC4E411ABE15 CRC64;

Query Match
Best Local Similarity 79.1%; Score 34; DB 12; Length 433;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAVVNC 7
DB 348 CSLPNC 354

RESULT 15
O8N5V5 PRELIMINARY: PRT: 446 AA.
ID O8N5V5:
AC O8N5V5:
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 1200003C23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RT Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC031567; AAH31567.1; -
DR InterPro: IPR000152; Asx_Hydroxyl.
DR InterPro: IPR006209; EGF_Like.
DR Pfam: PF00008; EGF_1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 1.
SQ SEQUENCE 446 AA; 47087 MW; F2CF5566AC083A6F CRC64;

Query Match
Best Local Similarity 79.1%; Score 34; DB 4; Length 446;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAVVNC 7
DB 11:111

```

Fri Sep 5 09:16:09 2003

us-09-580-893d-64.rspt

Db 212 CAELPNC 218

Search completed: September 4, 2003, 21:06:48
Job time : 28.815 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:15:56 ; Search time 41.1765 Seconds
(without alignments)
30.838 Million cell updates/sec

Title: US-09-580-893d-65

Perfect score: 49

Sequence: 1 CGAVPNC 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	49	100.0	8 23 AAE16727	Ligamentum nuchae
2	49	100.0	8 23 AAE16730	Ligamentum nuchae
3	43	87.8	8 23 AAE16716	Ligamentum nuchae
4	43	87.8	8 23 AAE16719	Ligamentum nuchae
5	43	87.8	91 23 ABP34373	Human ORE3346 prot
6	40	81.6	137 22 AAU42999	Propionibacterium
7	39	79.6	395 22 AAM25607	Human protein sequ
8	39	79.6	520 22 AAM23901	Rat EST encoded pr
9	39	79.6	584 22 AAB67465	Amino acid sequenc

10	39	79.6	594	22	AA675376	Human colon cancer
11	39	79.6	748	22	AA660097	Human transport pr
12	39	79.6	816	22	AA677820	Human ion channel
13	39	79.6	816	23	AB661532	Human transporter
14	37	75.5	283	23	AAE12888	Murine chordin pro
15	37	75.5	469	23	AB697384	Novel human protei
16	37	75.5	497	22	AAU16361	Human novel secret
17	37	75.5	497	24	ABU55430	Human novel polype
18	37	75.5	534	22	AAU15908	Human novel secret
19	37	75.5	534	24	ABU54977	Human novel polype
20	37	75.5	707	22	AAH83718	Human polypeptide,
21	37	75.5	707	22	AAH83234	Human protein sequ
22	37	75.5	801	22	AAE03738	Human chordin (CH
23	37	75.5	867	20	AAU01711	A human huchordin
24	37	75.5	867	20	ABG75647	Human TANGO-66/huc
25	37	75.5	915	23	ABG59661	Human secreted pro
26	37	75.5	915	23	ABG59661	Human secreted pro
27	37	75.5	936	22	AAE03735	Murine mature chor
28	37	75.5	948	22	AAE03736	Murine chordin (CH
29	37	75.5	948	23	AAU75168	Mouse chordin poly
30	37	75.5	954	19	AAU48978	Mature human chord
31	37	75.5	954	20	AAU17821	Human PRO243 prote
32	37	75.5	954	21	AAU01312	Human PRO243 polyp
33	37	75.5	954	23	AAU06132	Human PRO243 polyp
34	37	75.5	954	24	ABU67122	Human PRO polypept
35	37	75.5	954	24	ABU64917	Human secreted/tira
36	37	75.5	954	24	ABU60231	Human PRO polypept
37	37	75.5	954	24	ABU58351	Novel human secret
38	37	75.5	954	24	ABU55921	Human secreted/tira
39	37	75.5	954	24	ABU56302	Human secreted/tira
40	37	75.5	954	24	ABU57237	Human PRO243 prote
41	37	75.5	954	24	ABU11303	Human PRO243 prote
42	37	75.5	955	22	AAE12889	Human chordin prot
43	37	75.5	955	23	ABG31265	Human chordin (CH
44	37	75.5	962	22	AAE03734	Murine chordin (CH
45	36	73.5	29	13	AAE26998	Porcine IGFBP-6 tr

ALIGNMENTS

RESULT 1	
AAE16727	AAE16727 standard; peptide: 8 AA.
XX	
AC	AAE16727;
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPN derive peptide #7.
XX	
KW	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; candidant;
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypotensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW	restenosis post-angioplasty; antiangiinal; anticoagulant; thrombolytic.
XX	
OS	Ligamentum nuchae.
OS	Synthetic.
PN	WO200191700-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-0517384.
XX	
PR	30-MAY-2000; 2000US-0580110.
PR	30-MAY-2000; 2000US-0580156.
PR	30-MAY-2000; 2000US-0580893.
PR	30-MAY-2000; 2000US-0584001.
XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGINGING LLC.
XX	

PI	Melts TF	Sandberg LB, Jimenez F;
XX	DR	WPI: 2002-106259/14.
XX	PT	Composition for improving the elasticity of tissue comprises peptide
XX	PT	substances which mimic the sequences found in elastin
XX	PS	Claim 24; Page 21; 53pp; English.
XX	CC	The present invention relates to a composition for the treatment of
XX	CC	mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
XX	CC	their biological equivalents. The composition is useful not only for
XX	CC	improving the cosmetic appearance of skin, but for the treatment of
XX	CC	various skin disorders associated with reduced elastin formation and for
XX	CC	numerous other tissue disorders especially of blood vessels including
XX	CC	hypertension, coronary heart disease, arteriosclerosis, angina, coronary
XX	CC	thrombosis, chronic obstructive pulmonary disease and restenosis post-
XX	CC	angioplasty. They may also be useful in coating surgical appliances i.e.
XX	CC	stents. The use of peptides and peptide fragments which mimic those found
XX	CC	in elastin overcomes the problem of solubility of the dense cross-linked
XX	CC	structure of elastin itself, which renders it difficult to use.
XX	CC	therapeutically. The peptides are found to improve elastin production in
XX	CC	treated tissues. The present sequence is elastin peptide (HEP), VVPN derived peptide.
XX	Sequence	8 AA;
XX	Query Match	100.0%; Score 49; DB 23; Length 8;
XX	Best/Local Similarity	100.0%; Pred. No. 9.3e+05;
XX	Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	1 CGAVVPMC	8
Db	1 CGAVVPMC	8
RESULT 2		
AAE16730		
ID	AAE16730 standard; peptide; 8 AA.	
XX	AAE16730;	
XX	09-APR-2002 (first entry)	
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPN derived peptide #10.	
XX	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cartilage;	
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;	
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;	
KW	hypertensive; chronic obstructive pulmonary disease; antiarteriosclerotic;	
KW	restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;	
XX		
OS	Ligamentum nuchae.	
OS	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FT	Modified-site	1
FT		/note="This residue forms a cyclic structure with Cys
FT		at position 8 with copper as a chelating agent"
FT	Modified-site	8
FT		/note="This residue forms a cyclic structure with Cys
FT		at position 1 with copper as a chelating agent"
XX		
PN	W0200191700-A2.	
XX		
PD	06-DEC-2001.	
XX		
PF	30-MAY-2001; 2001WO-US17384.	
XX		
PR	30-MAY-2000; 2000US-0580110.	
PR	30-MAY-2000; 2000US-0580156.	
PR	30-MAY-2000; 2000US-0580893.	

PR	30-MAY-2000; 2000US-0584001.
XX	
PA	(CONN-) CONNECTIVE TISSUE IMAGING LLC.
XX	
PI	Milts TF, Sandberg LB, Jimenez F;
XX	
DR	WPI: 2002-106259/14.
XX	
PT	Composition for improving the elasticity of tissue comprises peptide
PT	substances which mimic the sequences found in elastin -
PS	Claim 24; Page 21; 53pp; English.
CC	
CC	The present invention relates to a composition for the treatment of
CC	mammalian tissue comprising peptides (hydrolysed elastin peptide, HEP) or
CC	their biological equivalents. The composition is useful not only for
CC	improving the cosmetic appearance of skin, but for the treatment of
CC	numerous other tissue disorders associated with reduced elastin formation and for
CC	hypertension, coronary heart disease, arteriosclerosis, angina, coronary
CC	thrombosis, chronic obstructive pulmonary disease and restenosis post-
CC	angioplasty. They may also be useful in creating surgical appliances i.e.
CC	stents. The use of peptides and peptide fragments which mimic those found
CC	in elastin overcomes the problem of solubility of the dense cross-linked
CC	structure of elastin itself, which renders it difficult to use
CC	therapeutically. The peptides are found to improve elastin production in
CC	treated tissues. The present sequence is Ligamentum nuchae hydrolysed
CC	elastin peptide (HEP), VVPM derived cyclic peptide.
XX	
SQ	Sequence 8 AA:
OY	Query Match 100.0%; Score 49; DB 23; Length 8;
DB	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 CGAVVPNC 8 1 CGAVVPNC 8
DB	
RESULT 3	
ID	AAE16716
XX	AAE16716 standard; peptide; 8 AA.
AC	AAE16716;
DT	09-APR-2002 (first entry)
DE	Ligamentum nuchae hydrolysed elastin peptide, VVPM derived peptide #10.
XX	
KW	Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
KW	vasotrophic; hypertension; skin disorder; tissue disorder; dermatological;
KW	coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
KW	hypertensive; chronic obstructive pulmonary disease; antiarteriosclerotic;
KW	restenosis post-angioplasty; antiaanginal; anticoagulant; thrombolytic;
cyclic.	
OS	Ligamentum nuchae.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Disulfide-bond 1..7
XX	
PV	WO200191700-A2.
XX	
PD	06-DEC-2001.
PF	30-MAY-2001; 2001WO-US17384.
PR	30-MAY-2000; 2000US-0580110.
PR	30-MAY-2000; 2000US-0580156.
PR	30-MAY-2000; 2000US-0580893.
PR	30-MAY-2000; 2000US-0584001.

XX (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PA Mlts TF, Sandberg LB, Jimenez F;
 PI WPI; 2002-106259/14.
 DR WPI; 2002-106259/14.
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 PS Claim 24; Page 20; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VWPQ derived cyclic peptide.
 XX Sequence 8 AA:
 SO
 Query Match 87.8%; Score 43; DB 23; Length 8;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAVVPMC 8
 Db 1 CGAVVPMC 8
 RESULT 4
 AAE16719
 ID AAE16719 standard; peptide; 8 AA.
 AC AAE16719;
 DT 09-APR-2002 (first entry)
 DE Ligamentum nuchae hydrolysed elastin peptide, VWPQ derived peptide #13.
 XX
 KW Hydrolysed elastin peptide; HEP; therapy; cosmetic appearance; cardiant;
 KW vasotropic; hypertension; skin disorder; tissue disorder; dermatological;
 KW coronary heart disease; arteriosclerosis; coronary thrombosis; angina;
 KW hypotensive; chronic obstructive pulmonary disease; antiatherosclerotic;
 KW restenosis post-angioplasty; antilanginal; anticoagulant; thrombolytic;
 KW cyclic.
 OS Ligamentum nuchae.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1 /note- "This residue forms a cyclic structure with Cys
 FT at position 8 with copper as a chelating agent"
 FT Modified-site 8 /note- "this residue forms a cyclic structure with Cys
 FT at position 1 with copper as a chelating agent"
 PN WO200191700-A2.
 PD 06-DEC-2001.
 PF 30-MAY-2001; 2001MO-US17384.
 XX

PR 30-MAY-2000; 2000US-0580110.
 PR 30-MAY-2000; 2000US-0580156.
 PR 30-MAY-2000; 2000US-0580893.
 PR 30-MAY-2000; 2000US-0584001.
 XX
 PA (CONN-) CONNECTIVE TISSUE IMAGING LLC.
 PI Mlts TF, Sandberg LB, Jimenez F;
 DR WPI; 2002-106259/14.
 XX Composition for improving the elasticity of tissue comprises peptide
 PT substances which mimic the sequences found in elastin -
 PS Claim 24; Page 20; 53pp; English.
 CC The present invention relates to a composition for the treatment of
 CC mammalian tissue comprises peptides (hydrolysed elastin peptide, HEP) or
 CC their biological equivalents. The composition is useful not only for
 CC improving the cosmetic appearance of skin, but for the treatment of
 CC various skin disorders associated with reduced elastin formation and for
 CC numerous other tissue disorders especially of blood vessels including
 CC hypertension, coronary heart disease, arteriosclerosis, angina, coronary
 CC thrombosis, chronic obstructive pulmonary disease and restenosis post-
 CC angioplasty. They may also be useful in coating surgical appliances i.e.
 CC stents. The use of peptides and peptide fragments which mimic those found
 CC in elastin overcomes the problem of solubility of the dense cross-linked
 CC structure of elastin itself, which renders it difficult to use
 CC therapeutically. The peptides are found to improve elastin production in
 CC treated tissues. The present sequence is Ligamentum nuchae hydrolysed
 CC elastin peptide (HEP), VWPQ derived cyclic peptide.
 XX Sequence 8 AA:
 SO
 Query Match 87.8%; Score 43; DB 23; Length 8;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAVVPMC 8
 Db 1 CGAVVPMC 8
 RESULT 5
 ABP34373
 ID ABP34373 standard; Protein; 91 AA.
 AC ABP34373;
 DT 08-JUL-2002 (first entry)
 DE Human ORF346 protein, SEQ ID NO:6692.
 XX
 KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; actinin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vunerary;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 OS Homo sapiens.
 OS
 PN WO200190366-A2.
 PD 29-NOV-2001.
 XX

FE		24 -MAY-2001; 2001IWO-US17076.
XX		
PR		24 -MAY-2000; 2000OUS-206690P.
XX		
PA	(GURA-) CURGAGEN CORP.	
XX		
PI	Leach MD, Shinkets RA;	
XX		
DR	WPI: 2002-106200/14.	
XX	N-PSDB: ABN78399.	
PT	Novel human polypeptides and polynucleotides useful for diagnosing,	
PT	preventing and treating cardiovascular disease, neurodegenerative,	
PT	hyperproliferative disorders and disorders related to organ	
XX	transplantation	
XX		
PS	Claim 10; Page 1920; 2508pp; English.	
XX		
CC	Sequences ABPJ1028-ABPJ5561 represent 4534 novel human proteins	
CC	designated ORF (open reading frame) 1-4534, and sequences ABN75054-	
CC	ABN79587 represent cDNAs encoding them. The invention also encompasses	
CC	polypeptides at least 80% identical to the ORF1-ORF4534 (collectively	
CC	referred to as ORFX) proteins, polynucleotides at least 85% identical to	
CC	the ORFX nucleic acid sequences, vectors and host cells comprising ORFX	
CC	polynucleotides, the recombinant production of ORFX proteins, antibodies	
CC	specific for ORFX proteins, methods of detecting ORFX polynucleotides and	
CC	polypeptides, methods of screening for modulators of ORFX expression or	
CC	activity, and methods of screening individuals for a predisposition to an	
CC	ORFX-associated disorder. The ORFX proteins of the invention have a wide	
CC	range of biological activities, such as cytokine, cell proliferation,	
CC	cell differentiation, immune modulation, hematopoiesis regulation,	
CC	tissue growth, angiogenesis, activin or inhibin activity, chemotactic/	
CC	chemokineic activity, haemostatic activity, thrombolytic activity,	
CC	receptor/ligand, antiinflammatory activity, tumour inhibition activity,	
CC	and antifibrotic activity, and may also be involved in the determination	
CC	of bodily characteristics, fertility and behaviour. ORFX proteins,	
CC	nucleic acids and antibodies may be used in the treatment of cancers,	
CC	other proliferative disorders such as psoriasis and benign tumours,	
CC	neurological disorders such as epilepsy and Alzheimer's disease,	
CC	cardiovascular diseases, immune system disorders, disorders related to	
CC	organ transplantation, disorders of tissue growth and regeneration,	
CC	diseases such as diabetes mellitus, hypothyroidism, and cholesterol	
CC	storage disease, and infectious diseases caused by viral, bacterial,	
CC	fungal and other pathogens. ORFX nucleic acids may also be used as a	
CC	source of primers and probes, in the detection of ORFX genomic sequences	
CC	of transcripts, in the identification and cloning of homologous	
CC	sequences, in genetic diagnosis, and in forensic biology. The ORFX	
CC	nucleic acids may additionally be used to produce transgenic animals	
CC	which may be useful for studying the function and/or activity of ORFX	
CC	protein, and in drug screening. The ORFX proteins may also be used as	
CC	immunogens to generate specific antibodies, which are useful in the	
CC	diagnosis, treatment and monitoring of ORFX-associated diseases.	
XX		
SO	Sequence	91 AA:
OY	Query Match	87.8%; Score 43; DB 23; Length 91;
Db	Best Local Similarity	Pred. No. 2.9;
	Matches 7; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
	1 CGAATPNC 8	
	38 CGAAVSPNC 45	
RESULT 6		
ID	AAU42999	
AC	AAU42999 standard; Protein; 137 AA.	
XX	AAU42999;	
DT	27-FEB-2002 (first entry)	
DE	Propionibacterium acnes immunogenic protein #3895.	

XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.
XX	
XX	Propionibacterium acnes.
XX	
XX	WO200181581-A2.
XX	
XX	01-NOV-2001.
XX	
XX	20-APR-2001; 2001WO-US12865.
XX	
XX	21-APR-2000; 2000US-199047P.
XX	02-JUN-2000; 2000US-208841P.
XX	07-JUL-2000; 2000US-216747P.
XX	
XX	(CORI-) CORIXA CORP.
XX	
XX	Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX	L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX	
XX	WPI: 2001-616774/71.
XX	
XX	N-PSDB; AAS59519.
XX	
XX	Propionibacterium acnes polypeptides and nucleic acids useful for treating acne vulgaris -
XX	
XX	Example 1: SEQ ID No 4194; 1069pp; English.
XX	
XX	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
XX	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 137 AA:
XX	
XX	Query Match 81.6%; Score 40; DB 22; Length 137;
XX	Best Local Similarity 87.5%; Pred. No. 14;
XX	Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX	
XX	1 CGAATPNC 8
XX	11111111
XX	45 CGAATPNC 52
XX	
XX	RESULT 7
XX	AAU25607
XX	AAU25607 standard; Protein: 395 AA.
XX	
XX	AAU25607;
XX	
XX	16-OCT-2001 (first entry)
XX	
XX	Human protein sequence SEQ ID NO:1122.
XX	

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antirheumatic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; vitruclide;
 KW anti-HIV; fungicide; antitumor; cardiovascular; antineoplastic; anemia;
 KW antiagregant; haemostatic; vulnary; antileuk; osteopathic; eczema;
 KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 XX Homo sapiens.
 OS
 XX W0200153455-A2.
 PN
 XX 26-JUL-2001.
 PD
 XX 22-DEC-2000; 2000WO-US35017.
 PF
 XX 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI: 2001-457603/49.
 DR N-PSDB; AAH99548.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 PS Claim 20; Page 232; 1217pp; English.
 CC AA999166 to AAH99904 encode the human proteins given in AA25225 to
 CC AA252963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; vitruclide; anti-HIV; fungicide; antitumor;
 CC cardiovascular; antineoplastic; antiallergic; haemostatic; vulnary;
 CC antileuk; osteopathic; dermatological; antisthmatic; antidiabetic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX
 SQ Sequence 395 AA;
 OY
 Query Match 79.6%; Score 39; DB 22; Length 395;
 Best Local Similarity 75.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 1 CGAVPNC 8
 169 CGIYPNC 176

RESULT 8
 ID AAM23901
 ID AAM23901 standard; Protein: 520 AA.
 AC AAM23901;
 DT 12-OCT-2001 (first entry)
 DE Rat EST encoded protein SEQ ID NO: 1426.
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Rattus norvegicus.
 XX
 XX W0200154477-A2.
 PN
 XX 02-AUG-2001.
 PD
 XX 25-JAN-2001; 2001WO-US02687.
 PF
 XX 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 PI
 XX WPI: 2001-476164/51.
 DR N-PSDB; AAH98560.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 20; Page 992-993; 1275pp; English.
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 SQ Sequence 520 AA;
 OY
 Query Match 79.6%; Score 39; DB 22; Length 520;
 Best Local Similarity 75.0%; Pred. No. 77;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 1 CGAVPNC 8
 358 CGIYPNC 365
 RESULT 9
 ID AAB67465
 ID AAB67465 standard; Protein: 584 AA.
 AC AAB67465;
 DT 15-MAY-2001 (first entry)
 DE Amino acid sequence of a calcium channel transport polypeptide.
 KW Calcium channel transport polypeptide; calcium trafficking;
 XX

KW neural disorder; HIV-induced dementia; immune system disorder;
 KW rheumatoid arthritis; muscular disorder; muscle contractile dysfunction;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; arrhythmia; renal disorder;
 KW proliferative disorder; cancer; lung carcinoma; breast cancer.
 XX
 OS Homo sapiens.
 PN W0200108635-A2.
 XX
 PD 08-FEB-2001.
 XX
 PF 27-JUL-2000; 2000WO-US20392.
 XX
 PR 28-JUL-1999; 99US-0145958.
 PR 18-AUG-1999; 99US-0149446.
 PR 14-MAR-2000; 2000US-0189064.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J, Shi Y;
 XX
 DR MPI: 2001-138604/14.
 DR N-PSDB: AAF55042.
 XX
 PT New isolated nucleic acid useful for diagnosing, detecting, or treating
 PT or preventing diseases associated with anomalies in calcium trafficking
 PT across the plasma membrane -
 XX
 PS Claim 11: Page 256-257; 259pp; English.
 XX
 CC The present sequence represents a calcium channel transport polypeptide.
 CC The polynucleotides, polypeptides, and antibodies are useful for
 CC preventing, treating, or ameliorating diseases associated with anomalies
 CC in calcium trafficking across the plasma membrane. They are used to
 CC diagnose, detect and treat or prevent diseases or conditions such as
 CC neural disorders (e.g. HIV-induced dementia), immune system disorders
 CC (e.g. rheumatoid arthritis), muscular disorders (e.g. muscle contractile
 CC dysfunction), reproductive disorders, gastrointestinal disorders,
 CC pulmonary disorders, cardiovascular disorders (e.g. arrhythmias), renal
 CC disorders, proliferative disorders, and/or cancerous diseases and
 CC conditions (e.g. lung carcinoma or breast cancer).
 CC
 SQ Sequence 584 AA;
 XX
 QY Query Match 79.6%; Score 39; DB 22; Length 584;
 DB Best Local Similarity 75.0%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CGAVVPMC 8
 358 CGIVFPMC 365
 DB
 RESULT 10
 AAG75376
 ID AAG75376 standard; Protein; 594 AA.
 XX
 AC AAG75376;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:6140.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN W0200122920-A2.
 XX
 PD 05-APR-2001.
 XX

PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR MPI: 2001-235357/24.
 DR N-PSDB: AAH34781.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11: Page 7588-7590; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB7789 represent sequences used in the exemplification of the
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 CC
 SQ Sequence 594 AA;
 XX
 QY Query Match 79.6%; Score 39; DB 22; Length 594;
 DB Best Local Similarity 75.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CGAVVPMC 8
 368 CGIVFPMC 375
 DB
 RESULT 11
 AAB60097
 ID AAB60097 standard; Protein; 748 AA.
 XX
 AC AAB60097;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human transport protein TPPF-17.
 XX
 KW Human; transport protein; TPPF; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W0200078953-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16668.
 XX
 PR 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.
 XX


```

PR 28-OCT-1999; 9905-0162287.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI; 2001-041424/05.
DR N-PSDB; AAF27717.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
XX Claim 2; Page 118-120; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TRPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
CC
XX
SQ Sequence 748 AA;
Query Match 79.6%; Score 39; DB 22; Length 748;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
   11 1111
Db 522 CGIVPNC 529

RESULT 12
AAG77820
ID AAG77820 standard; Protein; 816 AA.
XX
XX AAG77820;
AC
XX
XX 05-DEC-2001 (first entry)
DT
XX
DE Human ion channel 23927 protein (IC23927).
XX
XX IC23927; human; ion channel 23927; gene therapy; screening assay;
KW Predictive medicine; pharmacogenetics; pain disorder;
KW central nervous system disorder; ion transport; ion conductance;
KW membrane bound protein; membrane excitability; membrane polarisation;
KW synaptic transmission; pain signalling; cell activation; cell growth;
KW cell proliferation; cell differentiation; cell migration;
KW muscle contraction.
XX
XX Homo sapiens.
OS
XX
XX Key
FH
XX Location/Qualifiers
FT 114..128
FT Domain
FT /note="Transmembrane domain 1"
FT 146..168
FT Domain
FT /note="Transmembrane domain 2"
FT 178..195
FT Domain
FT /note="Transmembrane domain 3"
FT 199..210
FT Domain
FT /note="Transmembrane domain 4"
FT 233..254
FT Domain
FT /note="Transmembrane domain 5"
FT 269..287
FT Domain
FT /note="Pore domain: an overall hydrophobic domain"
FT 298..320
FT Domain
FT /note="Transmembrane domain 6"
FT 445..465
FT Domain
FT /note="Transmembrane domain 7"
FT 482..502
FT Domain
FT /note="Transmembrane domain 8"
FT 510..532
FT Domain

```

```

FT
FT Domain
FT /note="Transmembrane domain 9"
FT 539..554
FT /note="Transmembrane domain 10"
FT 570..594
FT Domain
FT /note="Transmembrane domain 11"
FT 637..653
FT Domain
FT /note="Pore domain: an overall hydrophobic domain"
FT 665..687
FT /note="Transmembrane domain 12"
XX
XX W0200164881-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06530.
XX
XX 29-FEB-2000; 2000US-0185938.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Curtis RAJ, Slios-santiago I;
XX
XX WPI; 2001-570694/64.
XX
XX N-PSDB; AAF78802, AAF78803.
XX
XX Isolated ion channel polypeptide IC23927 useful in screening assays and
XX treatment of disorders, e.g., central nervous system disorders and pain
XX disorders -
XX
XX Claim 14; Fig 1; 142pp; English.
XX
XX The present sequence represents the human ion channel 23927 (IC23927)
XX protein, which is claimed in the invention. The ion channel family of
XX proteins is a large family of membrane bound proteins responsible for a
XX wide range of transport and signalling functions in cells. The invention
XX comprises IC23927 protein and nucleic acids. The IC23927 nucleic acids
XX and proteins may be used in screening assays, predictive medicine (e.g.
XX diagnostic assays and pharmacogenetics) and treatment of disorders
XX characterised by insufficient production of IC23927 (e.g. central nervous
XX system disorders, pain disorders, or disorders of cellular growth
XX differentiation or migration). The IC23927 nucleic acids and proteins are
XX also useful as targets for developing modulating agents to regulate a
XX variety of cellular processes, such as: ion transport (e.g. ion
XX conductance); membrane excitability and/or polarisation; synaptic
XX transmission (e.g. pain signalling); cell activation; cell proliferation;
XX cell growth; cell differentiation; cell migration and muscle contraction;
XX Administration of a modulator of IC23927 (especially by gene therapy) may
XX be used to treat pain or a pain disorder.
XX
XX
XX Sequence 816 AA;
SQ
Query Match 79.6%; Score 39; DB 22; Length 816;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
   11 1111
Db 590 CGIVPNC 597

RESULT 13
ABG61532
ID ABG61532 standard; Protein; 816 AA.
XX
XX ABG61532;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX Human transporter and ion channel, TRICH2, Incyte ID 2907828CD1.
XX
XX Human; transporter and ion channel; TRICH; transport disorder;
XX neurological disorder; muscle disorder; immunological disorder; cancer;
XX scleroderma; systemic lupus erythematosus; allergy; leukaemia;

```

cell proliferative disorder; cervical cancer; breast cancer;
 neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
 myotonic dystrophy; cataplexy; endocrine disorder; diabetes;
 Grave's disease; gastrointestinal disorder; Crohn's disease;
 renal disorder; good pasture's syndrome; viral infection; cirrhosis;
 bacterial infection; fungal infection; parasitic infection;
 protozoal infection; helminthic infection; cardiovascular disorder;
 atherosclerosis; hepatic disease.

Homo sapiens.

WO200240541-A2.

23-MAY-2002.

25-OCT-2001; 2001WO-US46055.

27-OCT-2000; 2000US-243989P.
 03-NOV-2000; 2000US-245904P.
 09-NOV-2000; 2000US-247673P.
 17-NOV-2000; 2000US-249661P.
 20-NOV-2000; 2000US-252232P.
 01-DEC-2000; 2000US-250790P.

(INCY-) INCYTE GENOMICS INC.

Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;
 Walla MK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjaneva M;
 Ramkumar J, Arvizu C, Gietzen KJ, Lai PG, Azimzai Y, Khan FA;
 Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;
 Ison CH, Das D, Raumann BE, Policky JL, Kearney L;
 WPI: 2002-463570/49.
 N-PSDB; ABR83211.

New transporters and ion channels (TRICH) polypeptides, useful for
 diagnosing, preventing, and treating disorders associated with an
 abnormal expression or activity of TRICH, e.g. immunological, muscular
 or renal disorders

Claim 1; Page 132-134; 178pp; English.

The invention relates to human transporters and ion channels (TRICH)
 polypeptides, a naturally occurring amino acid sequence 90 % identical to
 TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 a recombinant polynucleotide comprising a promoter sequence operably
 linked to the TRICH polynucleotide, a cell transformed with the
 recombinant polynucleotide, a transgenic organism comprising the
 recombinant polynucleotide, an isolated antibody that binds specifically
 to TRICH, and screening for compounds which bind to TRICH, modulate
 TRICH, modulate TRICH expression or are anti/agonists of TRICH.
 The polypeptides are useful for diagnosing, treating, and
 preventing transport, neurological, muscle, immunological disorders
 (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 proliferative disorders such as cancers (e.g. leukaemia, cervical or
 breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 cataplexy), endocrine disorders (e.g. diabetes, Grave's disease),
 gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
 protozoal and helminthic infections, cardiovascular disorders (e.g.
 atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 other diseases and disorders detailed in the specification. They can also
 be used in assessing the effects of exogenous compounds on the
 expression of nucleic acid and amino acid sequences of transporters and
 ion channels. TRICH or its fragments may also be used in screening for
 compounds that specifically bind to and modulate the activity of TRICH.
 The polynucleotides can be used to create knock-in humanised animals or
 transgenic animals to model human disease. The present sequence
 represents a TRICH protein.

Sequence 816 AA;

Query Match 79.6%; Score 39; DB 23; Length 816;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVENC 8
 |||
 Db 590 CGIVFENC 597

RESULT 14
 AAE12888
 ID AAE12888 standard; Protein; 283 AA.

AC AAE12888;

DT 15-JAN-2002 (first entry)

DE Murine chordin protein.

KW Murine chordin-like-2 protein; CHL-2; hypotensive; cardiac; cytostatic;
 rheumatoid arthritis; cachexia; arrhythmia; osteoporosis; hepatitis;
 hypertension; myocardial infarction; stomach cancer; ulcer; gene therapy;
 antiinflammatory; tissue regeneration; osteoporosis; muscular dystrophy;
 congestive heart failure; hepatoma; angina.

XX Mus musculus.

PN WO200164885-A1.

PD 07-SEP-2001.

PE 02-MAR-2001; 2001WO-US06891.

PR 02-MAR-2000; 2000US-186462P.

PA (AMGE-) AMGEN INC.

PI Zhang K, Linh C, Nakayama N;

DR WPI: 2001-648246/74.

Novel Chordin-like-2 polypeptides useful for diagnosing and treating
 diseases and conditions affecting bone density, tissue regeneration,
 skeletal muscle, heart, stomach and liver, and for wound healing

Example 1; Fig 2; 167pp; English.

The invention relates to chordin-like-2 (CHL-2) proteins and their
 corresponding nucleic acid molecules. The invention also provides
 pharmaceutical compositions and methods for the diagnosis, treatment,
 amelioration, and/or prevention of diseases, disorders and conditions
 associated with CHL2 proteins. Chordin-like-2 (CHL-2) proteins are useful
 for diagnosing and treating diseases and conditions affecting bone
 density (osteoporosis, osteoporosis, osteoarthritis and rheumatoid
 arthritis), skeletal muscle (e.g., cachexia and muscular dystrophy),
 heart (arrhythmia, angina, hypertension, myocardial infarction and
 congestive heart failure), stomach (stomach cancer and ulcer), and
 liver (hepatitis and hepatoma), for promoting tissue regeneration
 and wound healing, for ex vivo expansion of haematopoietic stem cells
 and gene therapy performed through such cells. CHL2 protein is also
 useful for identifying CHL2 protein receptors. A selective binding
 agent (SBA) or fragment is useful for detecting or quantitating the
 amount of CHL2 protein. A transgenic non-human mammal is useful for
 determining whether a compound inhibits CHL2 protein activity or
 CHL2 protein production. The present sequence is murine chordin protein.

Sequence 283 AA;

Query Match 75.5%; Score 37; DB 22; Length 283;
 Best Local Similarity 75.0%; Pred. No. 98;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 ||| ||:|
 Db 232 CGAGVPHC 239

RESULT 15

ID ABB97384 standard; Protein; 469 AA.

AC ABB97384;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 652.

KM Human; antihaemic; valvular; antiinflammatory; immunomodulator;
 KW antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 expressed sequence tag.

OS Homo sapiens.

PN W0200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US26015.

PR 11-SEP-2000; 2000US-0659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABN32570.

PT An isolated polynucleotide for treating diseases associated with its
 encoded polypeptide such as cancer and multiple sclerosis -

PS Claim 20; SEQ ID NO 652; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.

SQ Sequence 469 AA;

Query Match 75.5%; Score 37; DB 23; Length 469;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVVPNC 8
 |:|||||

Db 389 GSVVPNC 395

Search completed: September 4, 2003, 21:00:36
 Job time : 41.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:54:37 ; Search time 12.4706 Seconds
(without alignments)
61.693 Million cell updates/sec

Title: US-09-580-893d-65

Perfect score: 49

Sequence: 1 CGAVVPNC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	77.6	167	1 AG0369	ferredoxin-type pr
2	37	75.5	469	2 T46929	hypothetical prote
3	37	75.5	469	2 T46930	hypothetical prote
4	36	73.5	61	2 AH2852	conserved hypothet
5	36	73.5	61	2 F97629	hypothetical prote
6	36	73.5	74	2 C23734	insulin-like growt
7	36	73.5	90	2 B86560	9 kDa-Cysteine-ric
8	36	73.5	90	2 A72064	cysteine rich oute
9	36	73.5	99	2 S47084	lipid transfer lik
10	36	73.5	105	1 FE1WT	ferredoxin [3Fe-4S
11	36	73.5	123	2 T52287	ferredoxin - Deino
12	36	73.5	240	2 A39842	insulin-like growt
13	36	73.5	837	2 F96561	unknown protein [1
14	35	71.4	109	2 F71644	ferredoxin (fdaA)
15	35	71.4	116	2 C97860	ferredoxin (fdaA)
16	35	71.4	118	1 PSNJ3M	phospholipase A2 (
17	35	71.4	118	1 PSNJ3M	phospholipase A2 (
18	35	71.4	250	2 G72495	probable polysufl
19	35	71.4	329	2 T17033	leucine rich repea
20	35	71.4	330	2 J02262	polygalacturonase
21	35	71.4	336	2 S15997	chitinase (EC 3.2.
22	35	71.4	341	2 T05764	hypothetical prote
23	35	71.4	506	2 T35923	probable tRNA synt
24	35	71.4	532	2 T47335	hypothetical prote
25	35	71.4	572	2 T27869	sphingomyelin phos
26	35	71.4	579	2 JC7629	membrane-type friz
27	35	71.4	605	2 T15291	sphingomyelin phos
28	35	71.4	612	2 C90374	hypothetical prote
29	35	71.4	653	2 T25194	hypothetical prote

30	34	69.4	92	2 T04395	probable phosphol
31	34	69.4	129	2 B35216	EP14 protein - fow
32	34	69.4	196	2 T46525	probable transcrip
33	34	69.4	368	2 H69335	iron-sulfur cluste
34	34	69.4	369	2 F69407	iron-sulfur cluste
35	34	69.4	893	2 T17276	hypothetical prote
36	34	69.4	921	2 S49965	probable membrane
37	34	69.4	1013	2 T10659	probable serine/th
38	34	69.4	1029	2 T00712	protein kinase hom
39	34	69.4	1188	2 D86236	protein ELAN23.5 l
40	34	69.4	1244	2 S37034	DNA-directed DNA p
41	33	67.3	81	2 D69477	hypothetical prote
42	33	67.3	102	2 S04126	probable phosphol
43	33	67.3	107	2 F82278	ferredoxin VC0804
44	33	67.3	226	2 JN0464	insulin-like growt
45	33	67.3	241	2 G87530	endonuclease III f

ALIGNMENTS

RESULT 1

AG0369 ferredoxin-type protein NapF [imported] - Yersinia pestis (strain C092)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 30-Sep-2002

C:Accession: AG0369

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittall, R.W.; Holden, M.T.G.; Prentice, M.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barril

Nature 413, 523-527, 2001

A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; PMID:21470413; PMID:11586360

A:Accession: AG0369

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-167 <NR>

A:Cross-references: GB:AL590842; PIDN:CAC92282.1; PID:q15980993; GSPDB:GN00175

C:Genetics:

A:Gene: napF

C:Superfamily: ferredoxin protein NapF; ferredoxin [2Fe-4S] homology

Query Match 77.6%; Score 38; DB 1; Length 167;

Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8

Db 146 CGACVPGC 153

RESULT 2

T46929 hypothetical protein DKFZp434E2220.1 - human

C:Species: Homo sapiens (man)

C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: T46929

R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Well, B.; Wiemann, S.

submitted to the Protein Sequence Database, February 2000

A:Reference number: 224136

A:Accession: T46929

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1469 <AA>

A:Cross-references: EMBL:AL157433

A:Experimental source: adult testis; clone DKFZp434E2220

C:Genetics:

A>Note: DKFZp434E2220.1

Query Match 75.5%; Score 37; DB 2; Length 469;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GAVVPNC 8
|:|||||
Db 389 GSVVPNC 395

RESULT 3
T46930
hypothetical protein DKFZp43400420.1 - human
C:Species: Homo sapiens (man)
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: T46930
R:Otterbein, B.; Obermaier, B.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224136
A:Accession: T46930
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-469 <AAA>
A:Cross-references: EMBL:AL157434
A:Experimental source: adult testis; clone DKFZp43400420
A:Note: DKFZp43400420.1

Query Match 75.5%; Score 37; DB 2; Length 469;
Best Local Similarity 85.7%; Pred. NO. 39;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

OY 2 GAVVPNC 8
|:|||||
Db 389 GSVVPNC 395

RESULT 4
AH2852
conserved hypothetical protein Atu2249 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AH2852
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, F.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AH2577; MUID:21608550; PMID:11743193
A:Accession: AH2852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-61 <KUR>
A:Cross-references: GB:AE008688; PID:AA143238.1; PID:917740722; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2249
A:Map position: circular chromosome

Query Match 73.5%; Score 36; DB 2; Length 61;
Best Local Similarity 62.5%; Pred. NO. 11;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
||:|||||
Db 6 CGTIVPGC 13

RESULT 5
F97629
hypothetical protein AGR_C_4091 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: F97629
R:Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F97629
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-61 <KUR>
A:Cross-references: GB:AE007869; PID:AAK87991.1; PID:915157403; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4091
A:Map position: circular chromosome

Query Match 73.5%; Score 36; DB 2; Length 61;
Best Local Similarity 62.5%; Pred. NO. 11;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
||:|||||
Db 6 CGTIVPGC 13

RESULT 6
C23734
insulin-like growth factor-binding protein 6 - pig (fragments)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 11-Jan-2000
C:Accession: C23734
R:Shimasaki, S.; Gao, L.; Shimonaka, M.; Ling, N.
Mol. Endocrinol. 5, 938-948, 1991
A:Title: Isolation and molecular cloning of insulin-like growth factor-binding protel
A:Reference number: A23734; MUID:92049376; PMID:1719383
A:Accession: C23734
A:Status: preliminary
A:Molecule type: Protein
A:Residues: 1-74 <SHI>
C:Superfamily: Insulin-like growth factor binding protein 1; thyroglobulin type I rep

Query Match 73.5%; Score 36; DB 2; Length 74;
Best Local Similarity 62.5%; Pred. NO. 13;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
||:|||||
Db 30 CGVTPNC 37

RESULT 7
B86560
9 kDa-Cysteine-rich lipoprotein [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: B86560
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86560
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <STO>
A:Cross-references: GB:BA000008; NID:98978929; PID:BA98764.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: omcA

Query Match 73.5%; Score 36; DB 2; Length 90;
Best Local Similarity 62.5%; Pred. NO. 15;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVVPNC 8
||:|||||
Db 55 CGSVVPSC 62

```

RESULT 8
A:2064
Cysteine rich outer membrane protein 3 CP0193 [imported] - Chlamydia pneumoniae (str
N:Alternate names: 9 kda-cysteine-rich lipoprotein
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
A:Accession: A72064; G81603
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: A72064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <ARN>
A:Cross-references: GB:AE001640; GB:AE001363; NID:g4376845; PIDN:AAD18698.1; PID:g437685
A:Experimental source: strain CW029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: G81603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <REA>
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38066.1; PID:g718912
C:Genetics:
A:Gene: omcA; CP0193

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 90;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVVPMC 8
DB 55 CGSVVPS 62

RESULT 9
S47084
Lipid transfer like protein - cowpae
C:Species: Vigna unguiculata (cowpae)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Sep-1997
C:Accession: S47084
R:Krause, A.; Stigrist, C.J.A.; Dehning, I.; Sommer, H.; Broughton, W.G.
submitted to the EMBL Data Library, June 1994
A:Description: Accumulation of transcript encoding a lipid transfer-like protein during
A:Reference number: S47084
A:Accession: S47084
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-99 <KRA>
A:Cross-references: EMBL:X79604; NID:g499033; PID:g499034

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 99;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAVVPMC 8
DB 92 CGVYPNC 99

RESULT 10
FERWT
ferredoxin [3Fe-4S] [4Fe-4S] - Thermus aquaticus (tentative sequence)
C:Species: Thermus aquaticus
C:Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 03-Nov-2000
C:Accession: A00216
R:Sato, S.; Nakazawa, K.; Hon-Nami, K.; Oshima, T.

```

```

Biochim. Biophys. Acta 668, 277-289, 1981
A:Title: Purification, some properties and amino acid sequence of Thermus thermophilu
A:Reference number: A90636; MUID:81184605; PMID:7225412
A:Accession: A00216
A:Molecule type: protein
A:Residues: 1-69; 97-105 <SAT>
A:Experimental source: strain HB8; ATCC 27634
R:Hille, R.; Yoshida, T.; Tarr, G.E.; Williams Jr., C.H.; Ludwig, M.I.; Fee, J.A.; Ke
J. Biol. Chem. 258, 13068-13073, 1983
A:Title: Studies of the ferredoxin from Thermus thermophilus.
A:Reference number: A92402; MUID:84032522; PMID:6313685
A:Contents: annotation; composition
A:Note: we have positioned residues 70-96 by homology with other Azotobacter-type fer
C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
C:Keywords: 3Fe-4S; 4Fe-4S; duplication; electron transfer; iron-sulfur protein; meta
F.1-57/Domain: ferredoxin 2[4Fe-4S] homology <FER>
F.16/49/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted
F.20/39/42/45/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match
Best Local Similarity 73.5%; Score 36; DB 1; Length 105;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPMC 8
DB 42 CGACVPAC 49

RESULT 11
D75287
ferredoxin - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: D75287
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <WHI>
A:Cross-references: GB:AE002064; GB:AE000513; NID:g6460134; PIDN:AAF11876.1; PID:g646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2330
A:Map position: 1
C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 123;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPMC 8
DB 88 CGACVPAC 95

RESULT 12
A39842
insulin-like growth factor-binding protein 6 precursor - human
N:Alternate names: IGFBP-6; insulin-like growth factor II-binding protein
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_revision 31-Dec-1995 #text_change 03-Dec-1999
C:Accession: A39842; B23734; PH0144; A35470; C35803
R:Kiefer, M.C.; Mastarz, F.R.; Bauer, D.M.; Zapf, J.
J. Biol. Chem. 266, 9043-9049, 1991
A:Title: Identification and molecular cloning of two new 30-kDa insulin-like growth f
A:Reference number: A39842; MUID:91225006; PMID:1709161
A:Accession: A39842
A:Molecule type: mRNA

```

A:Residues: 1-240 <KIE>
 A:Cross-references: GB:M62402; NID:9184813; PIDN:AA06187.1; PID:9184814
 R:Shimasaki, S.; Gao, L.; Shimonaka, M.; Ling, N.
 Mol. Endocrinol. 5, 938-948, 1991
 A:Title: Isolation and molecular cloning of insulin-like growth factor-binding protein-6
 A:Reference number: A23734; MUID:92049376; PMID:1719383
 A:Accession: B23734
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 'C', 3-240 <SHI>
 A:Cross-references: GB:M69054; NID:9183893; PIDN:AA08070.1; PID:9183894
 R:Andress, D.L.; Birnbaum, R.S.
 Biochem. Biophys. Res. Commun. 176, 213-218, 1991
 A:Title: A novel human insulin-like growth factor binding protein secreted by osteoblast
 A:Reference number: PH0143; MUID:91207395; PMID:1850257
 A:Accession: PH0144
 A:Molecule type: protein
 A:Experimental source: cell line V-2
 A:Residues: 'X', 30-31, 'P', 33-39, 'X', 41-42 <AND>
 R:Marlin, J.L.; Willetts, K.E.; Baxter, R.C.
 J. Biol. Chem. 265, 4124-4130, 1990
 A:Title: Purification and properties of a novel insulin-like growth factor-II binding pr
 A:Reference number: A35470; MUID:90154107; PMID:2154495
 A:Accession: A35470
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 28, 'A', 30-39 <MAR>
 R:Zapf, J.; Kiefer, M.; Merryweather, J.; Masiarz, F.; Bauer, D.; Born, W.; Fischer, J.A.
 J. Biol. Chem. 265, 14892-14898, 1990
 A:Title: Isolation from adult human serum of four insulin-like growth factor (IGF) bindi
 ic tumor hypoglycemia.
 A:Reference number: A35803; MUID:90368661; PMID:1697583
 A:Accession: C35803
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 'AA', 30-31, 'H', 33-54, 'QXG' <ZAP>
 C:Genetics:
 A:Gene: GDB:IGFBP6
 A:Cross-references: GDB:127456; OMIM:146735
 A:Map position: 12pter-12qter
 C:Superfamily: Insulin-like growth factor binding protein 1; thyroglobulin type I repeat
 C:Keywords: glycoprotein
 F:1-27/Domain: signal sequence #status predicted <STG>
 F:28-240/Product: insulin-like growth factor-binding protein 6 #status predicted <MAR>
 F:163-234/Domain: thyroglobulin type I repeat homology <THYI>
 F:229/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 240;
 Best Local Similarity 62.5%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 Db 71 CGVYTPNC 78

RESULT 13
 E96561
 unknown protein [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: P96561
 R:Phelios, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chih, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huttar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maitl, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: P96561
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-837 <STO>
 A:Cross-references: GB:AE005173; NID:96850339; PIDN:AAE29402.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: P915.1
 A:Map position: 1

Query Match 73.5%; Score 36; DB 2; Length 837;
 Best Local Similarity 62.5%; Pred. No. 96;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 Db 469 CNTVLPNC 476

RESULT 14
 F71644
 ferredoxin (fdxA) RP829 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: F71644
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: F71644
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-109 <AND>
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CAA15254.1; PID:9386
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: fdxA; RP829
 C:Superfamily: ferredoxin 214Fe-4S; ferredoxin 214Fe-4S homology
 F:2-58/Domain: ferredoxin 214Fe-4S homology <FER>

Query Match 71.4%; Score 35; DB 2; Length 109;
 Best Local Similarity 62.5%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 Db 43 CGVCVPDC 50

RESULT 15
 C97860
 ferredoxin [Imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: C97860
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: C97860
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-116 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AMU03821.1; PID:g15620421; GSPDB:GN00173
 C:Genetics:
 A:Gene: fdxA
 C:Superfamily: ferredoxin 214Fe-4S; ferredoxin 214Fe-4S homology

Query Match 71.4%; Score 35; DB 2; Length 116;
 Best Local Similarity 62.5%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 Db 11 11:1

Fri Sep 5 09:16:10 2003

us-09-580-893d-65.rpr

Page 5

Db 47 CGVCPDC 54

Search completed: September 4, 2003, 21:10:53
Job time : 12.4706 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2003, 20:18:11; Search time 6.94118 Seconds
(without alignments)
54.200 Million cell updates/sec

Title: US-09-580-893D-65
Perfect score: 49
Sequence: 1 CGAVVNC 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	37	75.5	948 1	CHRD_MOUSE
2	37	75.5	955 1	CHRD_HUMAN
3	36	73.5	78 1	FER_THET
4	36	73.5	90 1	OM3_CHLPN
5	36	73.5	99 1	NLTP_VIGUN
6	36	73.5	240 1	IBPE_HUMAN
7	35	71.4	108 1	FER_RICPR
8	35	71.4	118 1	PA23_NAJMO
9	35	71.4	118 1	PA23_NAJNG
10	35	71.4	330 1	PGI2_ARATH
11	35	71.4	330 1	PGI2_PYRGO
12	35	71.4	336 1	CHT2_ORYSA
13	35	71.4	485 1	MKR4_HUMAN
14	34	69.4	129 1	V061_F0MPV
15	34	69.4	715 1	SYM_METMA
16	34	69.4	921 1	YIB7_YEAST
17	34	69.4	1207 1	DPOL_ASFB7
18	34	69.4	1244 1	DPOL_ASFL6
19	33	67.3	68 1	NLW2_PEARU
20	33	67.3	102 1	NLW2_HORVU
21	33	67.3	152 1	PA23_ORHHA
22	33	67.3	152 1	PA2H_LATSE
23	33	67.3	152 1	PA2I_LATSE
24	33	67.3	226 1	IBPE_RAT
25	33	67.3	287 1	YJW1_ECOLI
26	33	67.3	304 1	IBP2_RAT
27	33	67.3	305 1	IBP2_MOUSE
28	33	67.3	486 1	EHSL_SCHPO
29	33	67.3	712 1	SYM_METAC
30	33	67.3	894 1	CYAS_HUMAN
31	32	65.3	55 1	FER_PEPAS
32	32	65.3	55 1	FER_CLOAC
33	32	65.3	55 1	FER_CLOSP

ALIGNMENTS

RESULT 1	ID	CHRD_MOUSE	STANDARD	PRT	948 AA
AC	0920E2				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Chordin precursor.				
GN	CHRD.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RP	SEQUENCE FROM N.A.				
RP	SRRAIN-B6SLI/FL;				
RA	La B., Bachiller D., Agius E., Piccolo S., De Robertis E.M.;				
RT	"Bmp-binding domains in the chordin secreted protein.";				
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99000848; PubMed=9782094;				
RA	Pappano W.N., Scott I.C., Clark T.G., Eddy R.L., Shows T.B.,				
RA	Greenspan D.S.;				
RT	"Coding sequence and expression patterns of mouse chordin and mapping				
RL	of the cognate mouse chrd and human CHRD genes.";				
CC	Genomics 52:236-239(1998).				
CC	-1- FUNCTION: Dorsalizing factor. Key developmental protein that				
CC	dorsalizes early vertebrate embryonic tissues by binding to				
CC	ventralizing TGF-beta family bone morphogenetic proteins (BMPs)				
CC	and sequestering them in latent complexes.				
CC	-1- SUBCELLULAR LOCATION: Secreted (By similarity).				
CC	-1- DEVELOPMENTAL STAGE: Detected at high levels of a in 7 dpc mouse				
CC	embryos; its level decrease at later developmental stages and in				
CC	adult tissues.				
CC	-1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.				
CC	-1- SIMILARITY: Contains 4 WFC domains.				
CC	-1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; AF096276; AAC1895.1; -				P80168 clostridium
DR	EMBL; AF069501; AAC6867.1; -				P00204 chlorobium
DR	MGI; MGI:1313268; Chrd.				O8K26 chlorobium
DR	GO; GO:0007420; P:brain development; IMP.				P00205 chlorobium
DR	GO; GO:0007389; P:pattern specification; IMP.				O75508 homo sapien
DR	GO; GO:0001501; P:skeletal development; IMP.				P09743 escherichia
DR	InterPro: IPR006559; SOG.				P54356 drosophila
DR	InterPro: IPR001007; VMF_C.				P14172 rhodospira
DR	Pfam: PF00093; vwc; 4.				P24200 escherichia
DR	SMART; SM00566; SOG; 3.				P24626 oryza sativ
					P35334 phaseolus v
					P04998 medicago sa

```

DR SMART: SM00214; VWC: 4.
DR PROSITE: PS01208; VWF_C_1; 2.
DR PROSITE: PS0184; VWF_C_2; 2.
KW Developmental protein; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 948 CHORDIN.
FT DOMAIN 49 126 VWC 1.
FT DOMAIN 699 759 VWC 2.
FT DOMAIN 779 845 VWC 3.
FT DOMAIN 867 927 VWC 4.
FT CARBOHD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 877 877 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 948 AA; 101512 MW; 4DC2A0AD19BD2147 CRC64;

Query Match 75.5%; Score 37; DB 1; Length 948;
Best Local Similarity 75.0%; Pred. NO. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0

OY 1 CGAAPPNC 8
Db 897 CGAGVPHC 904

RESULT 2
CHRD_HUMAN STANDARD; PRT; 955 AA.
ID CHRD_HUMAN Q9H2X0; Q95254; Q9H2D3; Q9H2W8; Q9H2W9; Q9P0Z2; Q9P0Z3; Q9P0Z4;
AC Q9P0Z5;
DC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chordin precursor.
GN CHRD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
MEDLINE=21365001; PubMed=11472837;
RA Millet C., Lemaire P., Orsetti B., Guglielmi P., Francois V.;
RT "The human chordin gene encodes several differentially expressed
spliced variants with distinct BMP opposing activities.",
RL M ch. Dev. 106:85-96(2001).
[2]
SEQUENCE OF 115-955 FROM N.A. (ISOFORM 5).
MEDLINE=99000848; PubMed=9782094;
RA Pappano W.N., Scott I.C., Clark T.G., Eddy R.L., Shows T.B.,
RA Greenspan D.S.;
RT "Coding sequence and expression patterns of mouse chordin and mapping
of the cognate mouse chrd and human CHRD genes.",
RL Genomics 52:236-239(1998).
[3]
SEQUENCE OF 51-125; 705-762; 784-850 AND 872-932 FROM N.A.
RA Lu B., Bachiller D., Agius E., Larrain J., Piccolo S., Nieters A.,
RA de Robertis E.M.;
RT "Bmp binding modules in the chordin patterning protein.",
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: Dorsalizing factor. Key developmental protein that
dorsalizes early vertebrate embryonic tissues by binding to
ventralizing T g-beta family bone morphogenetic proteins (BMPs)
and sequestering them in latent complexes (By similarity).
-1- SUBCELLULAR LOCATION: Secreted (By similarity).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=5;
Comment=Experimental confirmation may be lacking for some
isoforms;
Name=1;
CC IsoId=Q9H2X0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H2X0-2; Sequence=VSP_001069, VSP_001070;
CC Name=3;
CC

```

[illegible]

RESULT 3
 FER_THETH STANDARD: PRT: 78 AA.
 ID PER_THETH
 AC P03942;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ferredoxin.
 OS Thermus thermophilus.
 OC Bacteria: Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 ON NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE=81184605; PubMed=7225412;
 RA Sato S., Nakazawa K., Hon-Nami K., Oshima T.,
 RT Purification, some properties and amino-acid sequence of Thermus
 thermophilus HB8 ferredoxin."
 RL Biochim. Biophys. Acta 668:277-289(1981).
 RN [2]
 RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.64 ANGSTROMS).
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE=21537789; PubMed=11681700;
 RA Macedo-Ribeiro S., Martins B.M., Pereira P.J., Buse G., Huber R.,
 RA Soullame T.;
 RT "New insights into the thermostability of bacterial ferredoxins:
 RT high-resolution crystal structure of the seven-iron ferredoxin from
 RT Thermus thermophilus."
 RL J. Biol. Inorg. Chem. 6:663-674(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.,
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CML029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: THIS OUTER MEMBRANE PROTEIN IS ASSOCIATED WITH
 CC DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES
 CC (EBBS). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER
 CC ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor.
 CC -1- DEVELOPMENTAL STAGE: EXTRACELLULAR ELEMENTARY BODY STAGE
 CC SPECIFIC.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: AE001640; AAD18698.1; -
 DR EMBL: AE002180; AAF38066.1; -
 DR EMBL: AP002547; BAA98764.1; -
 DR PIR: A72064; A72064.
 DR PIR: B86560; B86560.
 DR TIGR: CP0193; -
 DR InterPro: IPR003517; Chlam_CMP3.
 DR Pfam: PF03503; Chlam_CMP3.1.
 DR PRINTS: PR01335; CHLAMIDIAM3.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 19
 FT CHAIN 20 90
 FT LIPID 20 20
 FT SEQUENCE 90 AA; 5803725248B62608 CRC64;
 SO SEQUENCE

Query Match 73.5%; Score 36; DB 1; Length 78;
 Best Local Similarity 75.0%; Pred. No. 3.3;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

AC Q92725; Q9J015;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich outer membrane protein 3 precursor (9 kDa-cysteine-rich
 DE lipoprotein) (9KD-CRP)
 GN OMCA OR CPN0558 OR CP0193.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria: Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Ulfteback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MORN and Chlamydia
 RT pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.,
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CML029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: THIS OUTER MEMBRANE PROTEIN IS ASSOCIATED WITH
 CC DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES
 CC (EBBS). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER
 CC ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor.
 CC -1- DEVELOPMENTAL STAGE: EXTRACELLULAR ELEMENTARY BODY STAGE
 CC SPECIFIC.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: AE001640; AAD18698.1; -
 DR EMBL: AE002180; AAF38066.1; -
 DR EMBL: AP002547; BAA98764.1; -
 DR PIR: A72064; A72064.
 DR PIR: B86560; B86560.
 DR TIGR: CP0193; -
 DR InterPro: IPR003517; Chlam_CMP3.
 DR Pfam: PF03503; Chlam_CMP3.1.
 DR PRINTS: PR01335; CHLAMIDIAM3.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 19
 FT CHAIN 20 90
 FT LIPID 20 20
 FT SEQUENCE 90 AA; 5803725248B62608 CRC64;
 SO SEQUENCE

Query Match 73.5%; Score 36; DB 1; Length 90;

Best Local Similarity 62.5%; Pred. No. 3.8;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
11:11:1
DB 55 CGSVYPC 62

RESULT 5
NLRP_VIGUN
ID NLRP_VIGUN STANDARD: PRT: 99 AA.
AC 043681:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable nonspecific lipid-transfer protein AKCS9 precursor (LTP).
OS *Vigna unguiculata* (Cowpea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_Taxid=3917;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Red Caloona; TISSUE=Root hairs;
RX MEDLINE=94281682; PubMed=8012050;
RA Krause A., Sigrist C.J.A., Denning I., Sommer H., Broughton W.J.;
RT "Accumulation of transcripts encoding a lipid transfer-like protein
RT during deformation of nodulation-competent *Vigna unguiculata* root
RT hairs".
RT Mol. Plant Microbe Interact. 7:411-418(1994).
CC - FUNCTION: Potential lipid transfer protein.
CC - TISSUE SPECIFICITY: MOST TISSUES EXCEPT NODULES.
CC - DEVELOPMENTAL STAGE: EXPRESSION CORRELATES WITH ROOT HAIR
CC DEFORMATION.
CC - SIMILARITY: Belongs to the plant LTP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: X79604; CA56113.1;
DR PIR: S47084; S47084.
DR InterPro: IPR003612; AAI.
DR Pfam: PF00234; tryp_alpha_aml1.
DR SMART: SM00499; AAI.1.
DR PROSITE: PS00597; PLANT_LTP; FALSE_NEG.
KW Lipid-binding; Transport; Signal.
FT SIGNAL 1
FT CHAIN 34
FT 34 99
FT PROBABLE NONSPECIFIC LIPID-TRANSFER
FT PROTEIN AKCS9.
FT DISULFID 34 68 BY SIMILARITY.
FT DISULFID 42 56 BY SIMILARITY.
FT DISULFID 57 92 BY SIMILARITY.
FT DISULFID 66 99 BY SIMILARITY.
SQ SEQUENCE 99 AA: 10449 MW; B52615DFAC30AC30 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 99;
Best Local Similarity 62.5%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
11:11:1
DB 92 CGSVYPC 99

RESULT 6
IBP6_HUMAN
ID IBP6_HUMAN STANDARD: PRT: 240 AA.

AC P24592; Q14492;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin-like growth factor binding protein 6 precursor (IGFBP-6)
DE (IBP-6) (IGF-binding protein 6).
GN IGFBP6 OR IBP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Osteosarcoma;
RX MEDLINE=91225006; PubMed=1709161;
RA Kiefer M.C., Maslarz F.R., Bauer D.M., Zapf J.;
RT "Identification and molecular cloning of two new 30-kDa insulin-like
RT growth factor binding proteins isolated from adult human serum".
RT J. Biol. Chem. 266:9043-9049(1991).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=99189142; PubMed=10087296;
RA Ehrenborg E., Zazzi H., Lagercrantz S., Granqvist M., Hillerbrand U.,
RA Allander S.V., Larsson C., Luthman H.;
RT "Characterization and chromosomal localization of the human insulin-
RT like growth factor-binding protein 6 gene".
RT Mamm. Genome 10:376-380(1999).
RN 13
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holys S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko J., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 14
RP SEQUENCE OF 2-240 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92049376; PubMed=1719383;
RA Shimazaki S., Gao L., Shimomura M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
RT protein-6".
RT Mol. Endocrinol. 5:938-948(1991).
RN 15
RP PARTIAL SEQUENCE OF 28-42.
RX MEDLINE=91207395; PubMed=1850257;
RA Andress D.L., Birnbam R.S.;
RT "A novel human insulin-like growth factor binding protein secreted by
RT osteoblast-like cells".
RT Biochem. Biophys. Res. Commun. 176:213-218(1991).
RN 16
RP PRELIMINARY SEQUENCE OF 28-42.
RC TISSUE=Cerebrospinal fluid;
RX MEDLINE=90005986; PubMed=2551732;
RA Roghani M., Hossienlopp P., Lepage P., Bailland A., Binoux M.;
RT "Isolation from human cerebrospinal fluid of a new insulin-like
RT growth factor-binding protein with a selective affinity for IGF-II".

Query Match 71.4%; Score 35; DB 1; Length 108;
 Best Local Similarity 62.5%; Pred. No. 6.9;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 DB 42 CGVCVPCD 49

RESULT 8

PA23_NAJMO STANDARD; PRT: 118 AA.

AC P00604; 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

OC Phospholipase A2 isozyme CM-III (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase).

OS Najia mossambica (Mozambique cobra).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosaunia; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Elapidae; Elapinae; Najia.

OX NCBI_TaxID=8644;

RN [1]

RP SEQUENCE.

RC TISSUE-Venom;

RA MEDLINE=77222126; PubMed=880314;

RT Joubert F.J.;

RT "Naja mossambica mossambica venom. Purification, some properties and the amino acid sequences of three phospholipases A (CM-I, CM-II and CM-III).";

RL Biochim. Biophys. Acta 493:216-227(1977).

CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides.

CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-acylglycerophosphocholine + a fatty acid anion.

CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY. GROUP I SUBFAMILY.

CC PIR: A00744; PSNM3M.

DR HSSP; P00598; IPOA.

DR InterPro: IPR001211; PhospholipaseA2.

DR Pfam: PF00068; phoslip; 1.

DR PRINTS: PR00389; PPHPLIPASEA2.

DR PRODOM: PD000303; PhospholipaseA2; 1.

DR SMART; SM00085; PA2c; 1.

DR PROSITE: PS00119; PA2_ASP; 1.

DR PROSITE: PS00118; PA2_HIS; 1.

KW Hydrolyase; Lipid degradation; Calcium; Multigene family.

FT ACT_SITE 47 47 BY SIMILARITY.

FT ACT_SITE 92 92 BY SIMILARITY.

FT DISULFID 11 70 BY SIMILARITY.

FT DISULFID 26 117 BY SIMILARITY.

FT DISULFID 28 44 BY SIMILARITY.

FT DISULFID 43 98 BY SIMILARITY.

FT DISULFID 50 91 BY SIMILARITY.

FT DISULFID 59 84 BY SIMILARITY.

FT DISULFID 77 89 BY SIMILARITY.

FT METAL 27 27 BY SIMILARITY.

FT METAL 29 29 BY SIMILARITY.

FT METAL 31 31 BY SIMILARITY.

FT METAL 48 48 BY SIMILARITY.

FT METAL 118 AA; 13302 MW; DCE2A6A2A4C11BC2 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 118;
 Best Local Similarity 75.0%; Pred. No. 7.4;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 DB 84 CGAVCNC 91

RESULT 9

PA23_NAJMO STANDARD; PRT: 118 AA.

AC P00605; 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

DE Phospholipase A2 isozyme III (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase).

OS Najia nigricollis (Black-necked cobra).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosaunia; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Elapidae; Elapinae; Najia.

OX NCBI_TaxID=8654;

RN [1]

RP SEQUENCE.

RC TISSUE-Venom;

RA Baker D.;

RT "Structure and function of snake venom toxins.";

RT (In) Walter R., Melethofer J. (eds.);

RL Peptides: chemistry, structure and function, pp.17-30, Ann Arbor Science, Ann Arbor (1975).

CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides.

CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-acylglycerophosphocholine + a fatty acid anion.

CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY. GROUP I SUBFAMILY.

CC HSSP; P00598; IPOA.

DR InterPro: IPR001211; PhospholipaseA2.

DR Pfam: PF00068; phoslip; 1.

DR PRINTS: PR00389; PPHPLIPASEA2.

DR PRODOM: PD000303; PhospholipaseA2; 1.

DR SMART; SM00085; PA2c; 1.

DR PROSITE: PS00119; PA2_ASP; 1.

DR PROSITE: PS00118; PA2_HIS; 1.

KW Hydrolyase; Lipid degradation; Calcium.

FT ACT_SITE 47 47 BY SIMILARITY.

FT ACT_SITE 92 92 BY SIMILARITY.

FT DISULFID 11 70 BY SIMILARITY.

FT DISULFID 26 117 BY SIMILARITY.

FT DISULFID 28 44 BY SIMILARITY.

FT DISULFID 43 98 BY SIMILARITY.

FT DISULFID 50 91 BY SIMILARITY.

FT DISULFID 59 84 BY SIMILARITY.

FT DISULFID 77 89 BY SIMILARITY.

FT METAL 27 27 BY SIMILARITY.

FT METAL 29 29 BY SIMILARITY.

FT METAL 31 31 BY SIMILARITY.

FT METAL 48 48 BY SIMILARITY.

FT METAL 118 AA; 13268 MW; F6F2ACABA4C11BCF CRC64;

Query Match 71.4%; Score 35; DB 1; Length 118;
 Best Local Similarity 75.0%; Pred. No. 7.4;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
 DB 84 CGAVCNC 91

RESULT 10

PG12 ARATH STANDARD: PRT; 330 AA.

AC Q9M5J8; Q9FL58; 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Polyalacturonase inhibitor 2 precursor (Polygalacturonase-inhibiting protein) (PGIP-2).

GN PGIP2 OR AT5G06870 OR MOJ9.4.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

ON NCBI_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=98344145; PubMed=9679202;

RA Kaneo T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence features of the regions of 1,381,565 bp covered by twenty one physically assigned P1 and TAC clones.";

RT DNA Res. 5:131-145(1998).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;

RC Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;

RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the SSP consortium (Salk/Stanford/PDEC)."

RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

RP [3]

RA Park B., Jin Y., Nam S., Kim H.;

RT "Arabidopsis thaliana polygalacturonase inhibiting protein 2 (PGIP2) gene.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an important factor for plant resistance to phytopathogenic fungi.

CC -1- SUBCELLULAR LOCATION: Cell-wall associated.

CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITING PROTEIN FAMILY.

CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB010697; BAB11145.1; -

DR EMBL; AY035121; AAK59626.1; -

DR EMBL; AF229250; AAF69828.1; -

DR InterPro: IPR001611; LRR.

DR InterPro: IPR007090; LRR_plant.

DR Pfam: PF00560; LRR_3

DR PRINTS; PR00019; LEURICHRPT.

DR Signal; Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.

KW SIGNAL 1

FT CHAIN 1 21 POTENTIAL.

FT REPEAT 118 141 LRR 1.

FT REPEAT 142 166 LRR 2.

FT REPEAT 167 192 LRR 3.

FT REPEAT 194 215 LRR 4.

FT REPEAT 261 285 LRR 5.

FT REPEAT 287 308 LRR 6.

FT DISULFID 25 55 BY SIMILARITY.

FT DISULFID 56 63 BY SIMILARITY.

FT DISULFID 298 320 BY SIMILARITY.

FT DISULFID 322 329 BY SIMILARITY.

FT CARBOHYD 106 N-LINKED (GLCNAC. :) (POTENTIAL).

FT CARBOHYD 120 120 N-LINKED (GLCNAC. :) (POTENTIAL).

FT CARBOHYD 130 130 N-LINKED (GLCNAC. :) (POTENTIAL).

FT CARBOHYD 291 291 N-LINKED (GLCNAC. :) (POTENTIAL).

SO SEQUENCE 330 AA; 37067 MW; 991IDF7015DDC280 CRC04;

Query Match 71.4%; Score 35; DB 1; Length 330;

Best local Similarity 62.5%; Pred. No. 18;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVPNC 8

DB 322 CGALPSC 329

RESULT 11

PGIP_PIRCO STANDARD: PRT; 330 AA.

AC Q05091;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Polyalacturonase inhibitor precursor (Polygalacturonase-inhibiting protein).

DE PGIP.

GN Pyrus communis (Pear).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Maloideae; Pyrus.

ON NCBI_TaxID=23211;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Bartlett;

RX MEDLINE=94151422; PubMed=8108494;

RA Scott H.U., Powell A.L., Damon S.E., Greve L.C., Bennett A.B., Labavitch J.M.;

RT "Molecular characterization of a polygalacturonase inhibitor from Pyrus communis L. cv Bartlett.";

RL Plant Physiol. 102:133-138(1993).

RN [2]

RP CHARACTERIZATION.

RC STRAIN=cv. Bartlett; TISSUE=fruit;

RA Abu-Gourh A.A., Greve L.C., Labavitch J.M.;

RT "Purification and partial characterization of 'Bartlett' pear polygalacturonase inhibitors.";

RL Physiol. Mol. Plant Pathol. 23:111-122(1983).

CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an important factor for plant resistance to phytopathogenic fungi.

CC -1- SUBUNIT: Homodimer.

CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITING PROTEIN FAMILY.

CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L09264; AAA33865.1; -

DR PIR; J02262; J02262

DR InterPro: IPR001611; LRR.

DR InterPro: IPR007090; LRR_plant.

DR Pfam: PF00560; LRR_3.

DR PRINTS; PR00019; LEURICHRPT.

DR Signal; Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.

KW SIGNAL 1 24 POTENTIAL.

FT CHAIN 1 24

FT REPEAT 118 141 LRR 1.

FT REPEAT 142 166 LRR 2.

FT REPEAT 167 192 LRR 3.

```

FT REPEAT 194 215 LRR 4.
FT REPEAT 217 237 LRR 5.
FT REPEAT 261 285 LRR 6.
FT REPEAT 287 309 LRR 7.
FT DISULFID 27 37 BY SIMILARITY.
FT DISULFID 58 65 BY SIMILARITY.
FT DISULFID 298 320 BY SIMILARITY.
FT DISULFID 322 329 BY SIMILARITY.
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 330 AA; 36504 MM; 36BA2BCF420393E3 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 330;
Best Local Similarity 62.5%; Pred. No. 18;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGAVPNC 8
DB 322 CGAPLPC 329

RESULT 12
ID CH12_ORYSA STANDARD; PRT: 336 AA.
AC P25765;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Basic endochitinase 2 precursor (EC 3.2.1.14).
GN RCH10
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91238706; PubMed=2034221;
RA Zhu Q., Lamb C.J.;
RT "Isolation and characterization of a rice gene encoding a basic
RT chitinase.";
RL Mol. Gen. Genet. 226:289-296(1991).
CC -1- FUNCTION: THIS PROTEIN FUNCTIONS AS A DEFENSE AGAINST CHITIN
CC CONTAINING FUNGAL PATHOGENS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS 1A WHICH IS SIMILAR TO
CC CLASS 1B IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL
CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL
CC HYDROLASES).
DR PIR: S15997; S15997.
DR HSSP: P23951; 2BA4.
DR Gremene; P25765;
DR InterPro: IPR001002; Chitin_binding_1.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; Chitin_bind_1; 1.
DR Pfam: PF00182; Glyco_hydro_1; 1.
DR PRINTS: PR00451; CHITINBINDING.
DR PRODOM: PD000609; Chitin_binding_1; 1.
DR PRODOM: PD054900; Glyco_hydro_19; 1.
DR SMART: SM00270; ChEBD1; 1.
DR PROSITE: PS00026; CHITIN_BINDING; 1.
DR PROSITE: PS00773; CHITINASE_19_1; 1.
DR PROSITE: PS00774; CHITINASE_19_2; 1.
KW Hydrolyase; Glycosidase; Chitin degradation; Chitin-binding; signal;
KW Multigene family.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 336 BASIC ENDOCHITINASE 2.

```

```

FT DOMAIN 22 64 CHITIN-BINDING (BY SIMILARITY).
FT DISULFID 24 39 BY SIMILARITY.
FT DISULFID 33 45 BY SIMILARITY.
FT DISULFID 38 52 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
FT DISULFID 99 161 BY SIMILARITY.
FT DISULFID 174 182 BY SIMILARITY.
FT DISULFID 281 313 BY SIMILARITY.
SO SEQUENCE 336 AA; 35565 MM; 7F4EC126265DEA84 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 336;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GAVPNC 8
DB 30 GAVCPNC 36

RESULT 13
ID MKR4_HUMAN STANDARD; PRT: 485 AA.
AC Q13434;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Makorin 4 (zinc finger protein 127-Xp) (ZNF127-Xp).
GN MKR4 OR ZNF127L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrich B.D., Longstreet M., Gustashaw K., Nicholls R.D.,
RA Willard H.F.;
RT "An X-linked homologue of the autosomal imprinted gene ZNF127
RT escapes X chromosome inactivation.";
RL submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 4 C3H1-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U41315; AAA99070.1;
DR Genew: HGNC:7115; MKR4.
DR InterPro: IPR000571; Znf_CCH.
DR InterPro: IPR001841; Znf_fing.
DR Pfam: PF00097; Zf-C3HC4; 1.
DR Pfam: PF00642; Zf-CCCH; 4.
DR SMART: SM00184; RING; 1.
DR SMART: SM00356; Znf_C3H1; 4.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS00089; ZF_RING_2; 1.
KW Zinc-finger; Repeat.
FT ZN_FING 96 114 C3H1-TYPE 1.
FT ZN_FING 125 143 C3H1-TYPE 2.
FT ZN_FING 249 267 C3H1-TYPE 3.
FT DOMAIN 271 298 MAKORIN-TYPE CYS-HIS.
FT ZN_FING 316 370 RING-TYPE.
FT ZN_FING 405 425 C3H1-TYPE 4.
SO SEQUENCE 485 AA; 52909 MM; AE28B962544CEFF CRC64;

Query Match 71.4%; Score 35; DB 1; Length 485;
Best Local Similarity 62.5%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      1 CGAVPNC 8
DB      214 CGRTVPC 221

RESULT 14
ID      V061_F0MPV STANDARD; PRT; 129 AA.
AC      P21972;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-AUG-1991 (Rel. 19, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
GN      Putative CC-type chemokine FPV061.
OS      FPV061 OR FP14.
OC      Fowlpox virus (FPV).
OC      Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OX      NCBI_TaxID=10261;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=FP-1;
RX      MEDLINE=90324937; PubMed=2165135;
RA      Tartaglia J., Winslow J., Goebel S., Johnson G.P., Taylor J.,
RA      Paolletti E.;
RT      Nucleotide sequence analysis of a 10.5 kbp HindIII fragment of
RT      fowlpox virus: relatedness to the central portion of the vaccinia
RT      virus HindIII D region."
RL      J. Gen. Virol. 71:1517-1524(1990).
RN      (2)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20193820; PubMed=10729156;
RA      Alonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT      "The genome of fowlpox virus."
RL      J. Virol. 74:3815-3831(2000).
CC      -1- SIMILARITY: LOW SIMILARITY TO THE INTERCRINE BETA FAMILY (SMALL
CC      CYTOKINE C-C) (CHEMOKINE CC).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X17202; CAA35065.1; -
DR      EMBL: AF198100; AAF44405.1; -
DR      PIR: B35216; B35216.
KW      Hypothetical protein; Cytokine.
SQ      SEQUENCE 129 AA; 14110 MW; E0C25E79875E40DB CRC64;

Query Match      69.4%; Score 34; DB 1; Length 129;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      2 GAVVPNC 8
DB      87 GATFPNC 93

RESULT 15
SYM_METMA
ID      SYM_METMA STANDARD; PRT; 715 AA.
AC      OBPV14;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE      (MetRS).
GN      METG OR MM0867.
OS      Methanosarcina mazei (Methanosarcina frisia).
OC      Archaea; Euryarchaeota; Euryarchaeota orders Incertae sedis;
OC      Methanosarcinales; Methanosarcinaceae; Methanosarcina.

```

```

OX      NCBI_TaxID=2209;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=Goel / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
RX      MEDLINE=22120827; PubMed=12125824;
RA      Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA      Martinez-Arias R., Henne A., Wierer A., Baumer S., Jacobi C.,
RA      Brueggemann H., Llenad T., Christmann A., Boemcke M., Steckel S.,
RA      Bhattacharyya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA      Fritz H.-J., Gottschalk G.;
RT      "The genome of Methanosarcina mazei: evidence for lateral gene
RT      transfer between Bacteria and Archaea."
RL      J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC      -1- FUNCTION: Is required not only for elongation of protein synthesis
CC      but also for the initiation of all mRNA translation through
CC      initiator tRNA(fmet) aminoacylation (By similarity).
CC      -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC      diphosphate + L-methionyl-tRNA(Met).
CC      -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC      -1- SUBUNIT: Homodimer (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC      Metg subfamily 1.
CC      -1- SIMILARITY: Contains 1 tRNA-binding domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AE013312; AAM30563.1; AL1_INT1.
DR      HAMAP: MF_00098; fused; 1; AL1_INT1.
DR      Interpro: IPR004495; MetG_Cterm.
DR      Interpro: IPR001412; tRNA-synt_1.
DR      Interpro: IPR002304; tRNA-synt_met.
DR      Interpro: IPR002547; tRNA_bind.
DR      Pfam: PF01588; tRNA_bind; 1.
DR      Pfam: PIRSF001528; MetRS_dimerising; 1.
DR      PRINTS: PRO1041; TRNASTYTHMET.
DR      TIGRFAMs: TIGR00398; metG; 1.
DR      TIGRFAMs: TIGR00399; metG_C_term; 1.
DR      PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR      PROSITE: PS50886; TRBD; 1.
KW      Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW      RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
FT      SITE 20 30 "HIGH" REGION.
FT      SITE 334 338 "KMSKS" REGION.
FT      DOMAIN 613 715 tRNA-BINDING.
FT      METAL 151 151 ZINC (BY SIMILARITY).
FT      METAL 154 154 ZINC (BY SIMILARITY).
FT      METAL 163 163 ZINC (BY SIMILARITY).
FT      METAL 167 167 ZINC (BY SIMILARITY).
FT      BINDING 337 337 ATP (BY SIMILARITY).
SQ      SEQUENCE 715 AA; 79964 MW; 926A7E57855DA9CD CRC64;

Query Match      69.4%; Score 34; DB 1; Length 715;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 CGAVPNC 8
DB      473 CGGVVNC 480

Search completed: September 4, 2003, 21:01:48
Job time : 7.9418 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 4, 2003, 20:49:47; Search time 32.8235 Seconds
(without alignments)
62.895 Million cell updates/sec

Title: US-09-580-893D-65
Perfect score: 49
Sequence: 1 CGAVV/PNC 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	87.8	532	17	08TNS9
2	39	79.6	362	16	09RD10
3	39	79.6	857	4	090101
4	38	77.6	89	10	041378
5	38	77.6	167	16	08XCF1
6	38	77.6	304	3	08XCF1
7	37	75.5	61	16	092N61
8	37	75.5	83	16	098HT2
9	37	75.5	413	4	08N2W7
10	37	75.5	469	4	09NS53
11	37	75.5	469	4	09NS53
12	37	75.5	707	4	08N2K5
13	37	75.5	707	4	096SK7
14	36	73.5	29	6	09TRY1
15	36	73.5	61	16	08UD78
16	36	73.5	90	2	046183

17	36	73.5	111	4	09H2B5	09h2b5 homo sapien
18	36	73.5	123	16	09RR29	09rr29 delnoccocu
19	36	73.5	390	17	08TWP9	08twp9 methanopyr
20	36	73.5	697	5	09GZ35	09gz35 trypanosoma
21	36	73.5	837	10	09M819	09m819 arabidopsis
22	36	73.5	872	10	08H140	08h140 arabidopsis
23	36	73.5	890	10	08VYS2	08vys2 arabidopsis
24	35	71.4	67	10	P82900	P82900 triticum a
25	35	71.4	67	10	P82901	P82901 triticum a
26	35	71.4	95	2	09ZAK1	09zak1 anabaena sp
27	35	71.4	96	10	09FEK9	09fek9 triticum t
28	35	71.4	100	10	08H8Y6	08hy6 oryza sativ
29	35	71.4	100	10	08H8Y4	08hy4 oryza sativ
30	35	71.4	101	10	08H4W8	08h4w8 oryza sativ
31	35	71.4	116	16	09ZG41	09zg41 rickettsia
32	35	71.4	134	5	09U7C7	09u7c7 dictyosteli
33	35	71.4	176	12	09DGW8	09dgw8 gallid her
34	35	71.4	232	16	0932N8	0932n8 staphylococ
35	35	71.4	250	17	09Y8M7	09y8m7 aeropyrum
36	35	71.4	253	10	09LUX1	09lux1 pyrus pyrif
37	35	71.4	327	10	038695	038695 actinidia
38	35	71.4	329	10	038738	038738 antirrhinu
39	35	71.4	330	10	P93270	P93270 malus dome
40	35	71.4	330	10	09LKO8	09lko8 prunus mah
41	35	71.4	330	10	022522	022522 prunus arm
42	35	71.4	341	10	081843	081843 arabidopsi
43	35	71.4	506	16	09ZBH7	09zbh7 streptomyce
44	35	71.4	532	10	09M190	09m190 arabidopsi
45	35	71.4	564	5	010916	010916 caenorhabdi

ALIGNMENTS

RESULT 1						
ID	08TNS9	PRELIMINARY;	PRT;	532	AA.	
AC	08TNS9;					
DT	01-JUN-2002 (TREMURel. 21, Created)					
DT	01-JUN-2002 (TREMURel. 21, Last sequence update)					
DT	01-JUN-2002 (TREMURel. 21, Last annotation update)					
DE	Cellulosomal protein.					
GN	MA2203.					
OS	Methanosarcina acetivorans.					
OC	Archaea: Euryarchaeota: Methanococci; Methanosarcinales;					
OC	Methanosarcinaceae; Methanosarcina.					
OX	NCBI_TaxID=2214;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-CZA / ATCC 35395 / DSM 2834;					
RX	MEDLINE=21929760; PubMed=11932238;					
RA	Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,					
RA	Fitzhugh W., Galvo S., Engels R., Smirnov S., Alnoor D., Brown A.,					
RA	Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,					
RA	Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,					
RA	Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,					
RA	Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,					
RA	Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,					
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,					
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,					
RA	Metcalf W.W., Birren B.;					
RT	"The genome of Methanosarcina acetivorans reveals extensive metabolic					
RT	and physiological diversity."					
RL	Genome Res. 12:532-542(2002).					
DR	EMBL; AE010906; AAW05599.1; -.					
SK	Complete proteome.					
SO	SEQUENCE 532 AA; 60558 MW; 3A12ADFA8D67BB43 CRC64;					
Query Match	87.8%;	Score 43;	DB 17;	Length 532;		
Best Local Similarity	87.5%;	Pred. No. 3.4;				
Matches	7;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

QY 1 CGAVPNC 8
DB 172 CGAVPNC 179

RESULT 2

Q99DI0 PRELIMINARY; PRT; 362 AA.
AC Q99DI0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Hypothetical protein SC00901.
GN SC00901 OR SC01.34C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
Thompson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939107; CAB62692.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 362 AA; 38401 MW; 9119D6FDDC3F34DF CRC64;

QY

Query Match 79.6%; Score 39; DB 16; Length 362;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 1 CGAVPNC 8
43 CGAVPNC 50

RESULT 3

Q9ULQ1 PRELIMINARY; PRT; 857 AA.
AC Q9ULQ1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE KIAA1169 protein (Fragment).
GN KIAA1169.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.,
RT "Characterization of cDNA clones selected by the Genemark analysis
from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
DR EMBL; AB032995; BAA86483.2;
DR InterPro: IPR001682; Ca/Na_pore.
DR InterPro: IPR002111; Cat_channel_TPL.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001220; Lectin_legb.

DR InterPro: IPR005820; M-channel_nlg.

DR Pfam: PF00520; Ion_trans_2.

DR PROSITE: PS00307; LECTIN_LEGUME_BETA_1.

KM Ionic channel; Transmembrane.

FT NON_TER 1
SQ SEQUENCE 857 AA; 98609 MW; 439B63104579C992 CRC64;

Query Match 79.6%; Score 39; DB 4; Length 857;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
DB 631 CGAVPNC 638

RESULT 4

Q41378 PRELIMINARY; PRT; 89 AA.
AC Q41378;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)
DE Lipid transfer protein (Fragment).
OS Senecio odoratus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC Senecioneae; Senecio.
OX NCBI_TaxID=34237;
RN [1]
RP SEQUENCE FROM N.A.
RA Pye J., Kolatukudy P.E.;
RT "Epidermis-specific transcripts including one that encodes a new class
of lipid transfer proteins in Kleintia odorata.";
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; L33791; AAA33933.1;
DR InterPro: IPR003612; AAI.
DR Pfam: PF00234; try_alpha_amy1; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 89 AA; 9218 MW; FC20EB9543516C6 CRC64;

Query Match 77.6%; Score 38; DB 10; Length 89;
Best Local Similarity 62.5%; Pred. No. 5.9;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
DB 82 CGAVPNC 89

RESULT 5

08ZCF1 PRELIMINARY; PRT; 167 AA.
AC 08ZCF1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Ferredoxin-type protein NapF.
GN NAPF OR YPO3040 OR Y1442.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.:
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague."
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Stretcher J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Stretley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
 RA Perry R.D.:
 RT "Genome sequence of *Yersinia pestis* KIM-5".
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL: AJ414155; CAC92282.1; -
 DR EMBL: AE013747; AM85014.1; -
 DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
 DR InterPro: IPR004496; NapF.
 DR Pfam: PF00037; fer4; 3.
 DR TIGRPFAMS: TIGR00402; napF; 1.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; 3.
 KW Complete proteome.
 SQ SEQUENCE 167 AA; 18075 MW; E2B5594D80E8CE6C CRC64;

Query Match 77.6%; Score 38; DB 16; Length 167;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 ||| |||
 Db 146 CGACVPGC 153

RESULT 6
 Q8XOF1 PRELIMINARY; PRT; 304 AA.
 AC Q8XOF1;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Related to putative nucleotide binding protein (NUPB).
 GN B14A6.050.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algen V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mees H.W., Mannhaup G.:
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF670007; CAD21307.1; -
 SQ SEQUENCE 304 AA; 32237 MW; 2108A23CER2B0E56 CRC64;

Query Match 77.6%; Score 38; DB 3; Length 304;
 Best Local Similarity 75.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 ||| |||
 Db 195 CGFVCPNC 202

RESULT 7
 Q92N61 PRELIMINARY; PRT; 61 AA.
 ID Q92N61
 AC Q92N61;

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein R02361.
 GN R02361 OR SMC0269.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandembol M., Weidner S., Galibert F.:
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AF591790; CAC46940.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 61 AA; 7006 MW; A1F9EF343054001 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 61;
 Best Local Similarity 62.5%; Pred. No. 6.5;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 ||: ||| |
 Db 6 CGSLVPGC 13

RESULT 8
 Q98HT2 PRELIMINARY; PRT; 83 AA.
 ID Q98HT2
 AC Q98HT2;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein msr2723.
 GN MSR2723.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAFF30309;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.:
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003000; BAB49784.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 83 AA; 9748 MW; DB612DDCF7DFDD51 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 83;
 Best Local Similarity 62.5%; Pred. No. 8.7;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVVPNC 8
 ||: ||| |
 Db 28 CGSLVPGC 35

RESULT 9

```

Q8N2M7
ID 08N2M7 PRELIMINARY; PRT; 413 AA.
AC 09NS53;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029879; AAH29879.1; -.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00093; VWC; 3.
DR SMART; SM00214; VWC; 3.
DR PROSITE; PS01208; VWF_C; 2.
KM Hypothetical protein.
FT NON_CODING
SQ SEQUENCE 413 AA; 43822 MW; FCED20584BE0E06 CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 413;
Pred. No. 40;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAVPNC 8
Db 360 CGAGVPHC 367

RESULT 10
ID 09NS53 PRELIMINARY; PRT; 469 AA.
AC 09NS53;
DT 01-OCT-2002 (TReMBLrel. 15, Created)
DT 01-OCT-2002 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN DKFZP34E2220.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oltensaeider B., Obermaier B., Mewes H.W., Weill B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ductum adenocarcinoma;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL157433; CAB75657.1; -.
DR EMBL; BC017704; AAH17704.1; -.
DR InterPro; IPR006568; PSP.
DR SMART; SM00581; PSP; 1.
KM Hypothetical protein.
SQ SEQUENCE 469 AA; 51129 MW; 891FCBBB70026DAC CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 469;
Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVPNC 8
Db 389 GSVVPHC 395

```

```

RESULT 11
ID 09NS52 PRELIMINARY; PRT; 469 AA.
AC 09NS52;
DT 01-OCT-2002 (TReMBLrel. 15, Created)
DT 01-OCT-2002 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN DKFZP3400420.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oltensaeider B., Obermaier B., Mewes H.W., Weill B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL157434; CAB75658.1; -.
DR InterPro; IPR006568; PSP.
DR SMART; SM00581; PSP; 1.
KM Hypothetical protein.
SQ SEQUENCE 469 AA; 51097 MW; 891FCBBB6C6FDB8B CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 469;
Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVPNC 8
Db 389 GSVVPHC 395

RESULT 12
ID 08N2K5 PRELIMINARY; PRT; 707 AA.
AC 08N2K5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90157.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hiro Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074638; BAC11105.1; -.
DR InterPro; IPR006568; PSP.
DR InterPro; IPR001878; znf_CCHC.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00581; PSP; 1.
DR SMART; SM00343; znf_C2HC; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
KM Hypothetical protein.
SQ SEQUENCE 707 AA; 78505 MW; 7BE6B30CF19B2773 CRC64;

Query Match
Best Local Similarity 75.5%; Score 37; DB 4; Length 707;
Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVPNC 8
Db 627 GSVVPHC 633

```


RESULT 13
096SK7

PRELIMINARY; PRT; 707 AA.

AC 096SK7; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirooka S., Ishi S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuko Y., Kanehori K.;
RT "NEO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK027702; BAB5308.1; -
DR InterPro: IPR006568; PSP.
DR Pfam: PF00098; zf-CCHC; 1.
DR PRINTS: PR00939; C2HC2NFTNGER.
DR SMART: SM00581; PSP; 1.
DR SMART: SM00343; Znf_C2HC; 1.
DR PROSITE: PS0158; ZF_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 78475 MW; EDAE954E42672P56 CRC64;

Query Match 75.5%; Score 37; DB 4; Length 707;
Best Local Similarity 85.7%; Pred. No. 66;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAVVENC 8
1:|||||
DB 627 GSVVENC 633

RESULT 14
09TRY1
AC 09TRY1; PRELIMINARY; PRT; 29 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Insulin-like growth factor-binding protein-6 (Fragment).
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]
RP SEQUENCE.
RX MEDLINE-92049376; PubMed-1719383;
RA Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of Insulin-like growth factor-binding
protein-6."
RL Mol. Endocrinol. 5:938-948(1991).
DR HSSP: P24593; 1BOE.
DR InterPro: IPR000867; Insl_gro_fac-pr.
DR Pfam: PF00219; IGFBP; 1.
FT NON_TER 1
NON_TER 1
SQ SEQUENCE 29 AA; 3129 MW; C976C5ADEDD051F6 CRC64;

Query Match 73.5%; Score 36; DB 6; Length 29;
Best Local Similarity 62.5%; Pred. No. 5.1;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAVVENC 8
1:|||||
DB 5 CGVTPNC 12

RESULT 15
08UD78
ID 08UD78; PRELIMINARY; PRT; 61 AA.
AC 08UD78;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu2249.
GN ATU2249 OR AGR_C_4091.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Gurolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houtmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Iappas C., Marzel B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL: AE009174; AAL43238.1; -
DR EMBL: AE008140; AAK87991.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 61 AA; 6966 MW; B72DF39E30D5A1E2 CRC64;

Query Match 73.5%; Score 36; DB 16; Length 61;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGAVVENC 8
1:|||||
DB 6 CGTIVPNC 13

Search completed: September 4, 2003, 21:06:48
Job time: 32.9144 secs

THIS PAGE BLANK (USPTO)